iSCSI SAN 配置指南

ESX 4.1

ESXi 4.1

vCenter Server 4.1

在本文档被更新的版本替代之前，本文档支持列出的每个产品的版本和所有后续版本。要查看本文档的更新版本，请访问 http://www.vmware.com/cn/support/pubs。

ZH_CN-000288-00

vmware®

最新的技术文档可以从 VMware 网站下载：

http://www.vmware.com/cn/support/pubs/

VMware 网站还提供最近的产品更新信息。

您如果对本文档有任何意见或建议，请把反馈信息提交至：

docfeedback@vmware.com

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

北京办公室
北京市海淀区科学院南路 2 号
融科资讯中心 C 座南 8 层
www.vmware.com/cn

上海办公室
上海市浦东新区浦东南路 999 号
新梅联合广场 23 楼
www.vmware.com/cn

广州办公室
广州市天河北路 233 号
中信广场 7401 室
www.vmware.com/cn

目录

关于本文档

《iSCSI SAN 配置指南》说明如何将 VMware® ESX™ 和 VMware ESXi 系统与 iSCSI 存储区域网络 (SAN) 配合使用。本手册包括概念性的背景信息以及 ESX、ESXi 和 VMware vCenter™ Server 的安装要求。

目标读者

本手册的目标读者为熟悉虚拟机技术数据中心操作，且具有丰富经验的 Windows 或 Linux 系统管理员。

VMware 技术出版物词汇表

VMware 技术出版物提供了一个词汇表，其中包含一些您可能不熟悉的术语。有关 VMware 技术文档中所使用的术语的定义，请访问 http://www.vmware.com/support/pubs。

文档反馈

VMware 欢迎您提出宝贵建议，以便改进我们的文档。如有意见，请将反馈发送到 docfeedback@vmware.com。

VMware vSphere 文档

VMware vSphere 文档包括 VMware vCenter Server 和 ESX/ESXi 文档集。

技术支持和教育资源

您可以获取以下技术支持资源。有关本文档和其他文档的最新版本，请访问：http://www.vmware.com/support/pubs。

在线支持和电话支持

要通过在线支持提交技术支持请求、查看产品和合同信息以及注册您的产品，请访问 http://www.vmware.com/support。

客户只要拥有相应的支持合同，就可以通过电话支持，尽快获得对优先级高的问题的答复。请访问 http://www.vmware.com/support/phone_support.html。

支持服务项目

要了解 VMware 支持服务项目如何帮助您满足业务需求，请访问 http://www.vmware.com/support/services。

VMware 专业服务

VMware 教育服务课程提供了大量实践操作环境、案例研究示例，以及用作作业参考工具的课程材料。这些课程可以通过现场指导、教室授课的方式学习，也可以通过在线直播的方式学习。关于现场试点项目及实施的最佳实践，VMware 咨询服务可提供多种服务，协助您评估、计划、构建和管理虚拟环境。要了解有关教育课程、认证计划和咨询服务的信息，请访问 http://www.vmware.com/services。

将 ESX/ESXi 与 iSCSI 存储区域网络配合使用 1

您可以将 ESX/ESXi 与存储区域网络 (SAN) 配合使用，SAN 是将计算机系统连接至高性能存储子系统的专用高速网络。将 ESX/ESXi 与 SAN 配合使用可提供存储器整合，提高可靠性，并在灾难恢复方面提供帮助。

要将 ESX/ESXi 与 SAN 有效配合使用，您必须掌握 ESX/ESXi 系统与 SAN 概念的相关应用知识。此外，将 ESX/ESXi 主机设置为使用 Internet SCSI (iSCSI) SAN 存储系统时，必须了解存在的若干特殊注意事项。

本章讨论了以下主题：

了解虚拟化

VMware 虚拟化层在各款 VMware 桌面产品（如 VMware Workstation）及服务器产品（如 VMware ESX/ESXi）之间是通用的。该层为应用程序工作负载的开发、测试、交付及支持提供了一致平台。

虚拟化层构成方式如下：

- 每台虚拟机都运行各自的操作系统（客户机操作系统）和应用程序。
- 虚拟化层提供映射至特定物理设备份额的虚拟设备。这些设备包括虚拟化的 CPU、内存、I/O 总线、网络接口、存储适配器和设备、人机接口设备及 BIOS。

网络虚拟化

虚拟化层确保每台虚拟机与其他虚拟机相隔离。虚拟机仅可通过用于连接独立物理机的类似网络机制相互通信。

这种隔离方式使管理员可以构建内部防火墙或其他网络隔离环境，从而使得一些虚拟机可与外部连接，而另一些虚拟机只能通过虚拟网络与其他虚拟机连接。

存储器虚拟化

ESX/ESXi 提供主机级别的存储器虚拟化，即采用逻辑方式从虚拟机中抽象物理存储器层。

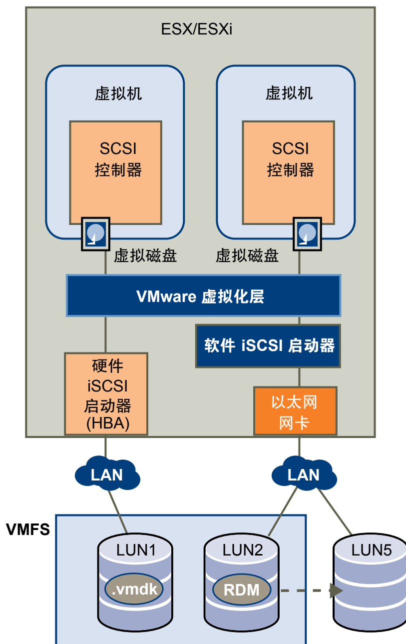
ESX/ESXi 虚拟机使用虚拟磁盘来存储其操作系统、程序文件，以及与其活动相关联的其他数据。虚拟磁盘是一个大型物理文件或一组文件，可以像任何其他文件一样轻松地对其进行复制、移动、存档和备份。您可以配置具有多个虚拟磁盘的虚拟机。

要访问虚拟磁盘，虚拟机需使用虚拟 SCSI 控制器。这些虚拟控制器包括 BusLogic 并行、LSI Logic 并行、LSI Logic SAS 和 VMware 准虚拟。虚拟机只能查看和访问以上类型的 SCSI 控制器。

虚拟机可通过上述某一虚拟 SCSI 控制器访问的每个虚拟磁盘，位于 VMware 虚拟机文件系统 (VMFS) 数据存储上、基于 NFS 的数据存储上或裸磁盘上。从虚拟机的角度而言，每个虚拟磁盘看上去都好像是与 SCSI 控制器连接的 SCSI 驱动器。实际的物理磁盘设备是通过并行 SCSI、iSCSI、网络还是主机上的光纤通道适配器来访问，这对客户机操作系统以及虚拟机上运行的应用程序而言是透明的。

图 1–1 概述了存储器虚拟化。图中所示为使用 VMFS 的存储器和使用裸设备映射的存储器。该图同时演示了如何通过 iSCSI HBA 或使用 iSCSI 启动器软件的通用网卡来访问 iSCSI 存储器。

图 1–1 iSCSI SAN 存储器虚拟化



iSCSI SAN 概念

如果您是 ESX/ESXi 管理员，并计划设置 ESX/ESXi 主机与 SAN 配合使用，那么您必须掌握 SAN 概念的相关应用知识。在某些出版资料中或 Internet 上可以找到 SAN 的相关信息。

iSCSI SAN 在计算机系统或主机服务器与高性能存储子系统之间使用以太网连接。SAN 组件包括主机服务器中的 iSCSI 主机总线适配器 (HBA) 或网络接口卡 (NIC)、传输存储流量的交换机和路由器、电缆、存储处理器 (SP) 及存储磁盘系统。

iSCSI SAN 使用客户端-服务器架构。客户端（也称为 iSCSI 启动器）在主机上操作。它通过发出 SCSI 命令并将封装到 iSCSI 协议中的这些命令传输到服务器来启动 iSCSI 会话。服务器也称为 iSCSI 目标。iSCSI 目标表示网络上的物理存储系统。它也可以由虚拟 iSCSI SAN（例如，正在虚拟机中运行的 iSCSI 目标仿真器）提供。iSCSI 目标通过传输必需的 iSCSI 数据响应启动器的命令。

iSCSI SAN 中的端口

iSCSI SAN 上单个可发现的实体（如启动器或目标）表示一个 iSCSI 节点。每个节点都有一个或多个将其连接到 SAN 的端口。

iSCSI 端口是 iSCSI 会话的端点。每个端口可通过多种方式进行标识。

IP 地址

每个 iSCSI 节点都可具有一个与其相关联的 IP 地址，以便网络上的路由和交换设备可以在服务器与存储器之间建立连接。此地址就像为了访问公司的网络或 Internet 而分配给计算机的 IP 地址一样。

iSCSI 名称

用于标识节点的全球唯一名称。iSCSI 使用 iSCSI 限定名 (IQN)、扩展唯一标识符 (EUI) 和网络地址授权 (NAA) 格式。

默认情况下，ESX/ESXi 为 iSCSI 启动器生成唯一 iSCSI 名称，例如，`iqn.1998-01.com.vmware:iscsitestox-68158ef2`。通常，无需更改默认值；但如果这样做，请确保输入的新 iSCSI 名称是全球唯一的。

iSCSI 别名

即一种替代 iSCSI 名称使用的更易于管理的 iSCSI 设备或端口名称。iSCSI 别名不是唯一的，它只是一个与端口关联的友好名称。

iSCSI 命名约定

iSCSI 使用一种特殊、唯一的名称来标识 iSCSI 节点（目标或启动器）。此名称类似于与光纤通道设备相关联的全球名称 (WWN)，可作为一种通用的节点识别方式使用。

iSCSI 名称通过两种不同方式格式化。最常见的是 IQN 格式。

有关 iSCSI 命名要求和字符串配置文件的更多详细信息，请参见 IETF 网站上的 RFC 3721、RFC 3722 和 RFC 3980。

iSCSI 限定名 (IQN) 格式

IQN 格式采用 `iqn.yyyy-mm.naming-authority:unique name` 的形式，其中：

- *yyyy-mm* 是命名机构成立的年份和月份。
- *naming-authority* 通常是命名机构的 Internet 域名的反向语法。例如，iscsi.vmware.com 命名机构的 iSCSI 限定名形式可能是 iqn.1998-01.com.vmware.iscsi。此名称表示 vmware.com 域名于 1998 年 1 月注册，iscsi 是一个由 vmware.com 维护的子域。
- *unique name* 是希望使用的任何名称，如主机的名称。命名机构必须确保在冒号后面分配的任何名称都是唯一的，例如：
  - iqn.1998-01.com.vmware.iscsi:name1
  - iqn.1998-01.com.vmware.iscsi:name2
  - iqn.1998-01.com.vmware.iscsi:name999

企业唯一标识符 (EUI) 格式

EUI 格式采用 `eui.`*`16 hex digits`* 的形式。

例如，`eui.0123456789ABCDEF`。

16 位十六进制数字是 IEEE EUI（扩展唯一标识符）格式的 64 位数的文本表示形式。前 24 位是 IEEE 向特定公司注册的公司 ID。后 40 位由持有该公司 ID 的实体分配，并且必须是唯一的。

iSCSI 启动器

要访问 iSCSI 目标，主机需要使用 iSCSI 启动器。启动器在主机与 iSCSI 目标之间传输封装到 iSCSI 协议中的 SCSI 请求和响应。

VMware 支持各种类型的启动器。

软件 iSCSI 适配器

软件 iSCSI 适配器是 VMkernel 中内置的 VMware 代码。借助于该适配器，主机可通过标准网络适配器连接到 iSCSI 存储设备。软件 iSCSI 适配器可在与网络适配器进行通信的过程中负责 iSCSI 处理过程。借助软件 iSCSI 适配器，您无需购买专用硬件便可使用 iSCSI 技术。

硬件 iSCSI 适配器

硬件 iSCSI 适配器是第三方适配器，可从主机卸载 iSCSI 和网络处理。硬件 iSCSI 适配器分为几种类别。

从属硬件 iSCSI 适配器

基于 VMware 网络以及由 VMware 提供的 iSCSI 配置和管理界面。

这种类型的适配器可以是一张为同一端口提供标准网络适配器和 iSCSI 卸载功能的卡。iSCSI 卸载功能取决于主机获取 IP、MAC 和其他用于 iSCSI 会话的参数的网络配置。iSCSI 许可的 Broadcom 5709 网卡就是一种从属适配器。

独立硬件 iSCSI 适配器

执行自己的网络及 iSCSI 配置和管理界面。

只提供 iSCSI 卸载功能，或者提供 iSCSI 卸载功能和标准网卡功能的卡就是一种独立硬件 iSCSI 适配器。iSCSI 卸载功能具有独立的配置管理，可分配 IP、MAC 和其他用于 iSCSI 会话的参数。QLogic QLA4052 适配器就是一种独立适配器。

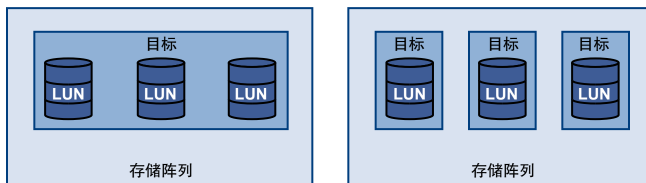
硬件 iSCSI 适配器可能需要获得许可。否则，将不会出现在 vSphere Client 或 vSphere CLI 中。有关许可信息，请联系供应商。

目标与 LUN 表示形式比较

在 ESX/ESXi 环境中，术语“目标”标识主机可以访问的单个存储单元。术语“存储设备”和“LUN”描述代表目标上的存储空间的逻辑卷。通常，术语“设备”和“LUN”在 ESX/ESXi 环境中表示通过存储器目标向主机显示的 SCSI 卷，并且该卷可以格式化。

不同的 iSCSI 存储供应商通过不同方式向服务器呈现存储器。有些供应商在单个目标上呈现多个 LUN，而有些供应商则向多个目标各呈现一个 LUN。虽然 ESX/ESXi 使用存储器的方式相似，但通过管理工具呈现信息的方式却不相同。

图 1–2 目标与 LUN 表示形式比较



这些配置中的每一个配置都有三个 LUN 可用。在第一种情况中，ESX/ESXi 可以检测到一个目标，但该目标具有三个可用 LUN。每个 LUN 均表示单独的存储卷。在第二种情况中，ESX/ESXi 可以检测到三个不同目标，每个目标都拥有一个 LUN。

基于 ESX/ESXi 的 iSCSI 启动器建立与各目标的连接。具有包含多个 LUN 的单个目标的存储系统通过一个连接向所有 LUN 输送流量。在具有三个目标（每个均有一个 LUN）的系统中，主机分别连接至三个 LUN。当您尝试从具有多个 iSCSI HBA 的 ESX/ESXi 主机（其中一个目标的流量可以设置为一个特定的 HBA，而另一个目标的流量可以使用其他 HBA）聚合多个连接上的存储流量时，此信息很有用。

多路径和路径故障切换

在主机服务器与存储器之间传输数据时，SAN 将使用一种叫做多路径的技术。使用多路径，您可以通过多条物理路径从 ESX/ESXi 主机到达存储系统上的 LUN。

通常，从主机到 LUN 的单一路径包括 iSCSI 适配器或网卡、交换机端口、连接电缆和存储控制器端口。如果路径中的任何组件出现故障，主机将选择另一条可用路径用于 I/O。检测故障路径并切换到另一条路径的过程称为路径故障切换。

存储系统类型

ESX/ESXi 可支持不同的存储系统和阵列。

主机支持的存储类型包括主动-主动、主动-被动和 ALUA 合规。

主动-主动存储系统
允许通过所有可用的存储器端口同时访问多个 LUN，而不会明显降低性能。所有路径始终保持活动状态，除非路径发生故障。

主动-被动存储系统
存储处理器在其中主动提供对给定 LUN 的访问权限的系统。其他处理器用作备用 LUN，并且可以主动提供对其他 LUN I/O 的访问权限。只能向指定 LUN 的活动端口成功发送 I/O。如果通过主动存储器端口访问失败，则其中一个被动存储处理器可通过服务器对其进行访问而激活。

非对称存储系统
支持非对称逻辑单元访问 (ALUA)。ALUA 合规存储系统对每个端口提供不同级别的访问。ALUA 允许主机确定目标端口的状态，并区分路径的优先级。主机会将某些活动路径用作主路径，其他用作辅助路径。

虚拟端口存储系统
允许通过一个虚拟端口访问所有可用的 LUN。这些是主动-主动存储设备，但通过单个端口隐藏其多个连接。ESX/ESXi 多路径无法检测与存储器的多个连接。这些存储系统以透明的方式处理端口故障切换和连接平衡。这通常称为透明故障切换。

发现、身份验证和访问控制

可以使用多种机制发现存储器并限制对它的访问。

必须配置主机和 iSCSI 存储系统以支持存储访问控制策略。

发现

发现会话是 iSCSI 协议的一部分，它返回可在 iSCSI 存储系统上访问的一组目标。ESX/ESXi 上提供的两类发现为：动态和静态。动态发现从 iSCSI 存储系统获取可访问目标的列表，而静态发现仅可按目标名称尝试访问一个特定目标。

身份验证

iSCSI 存储系统通过名称和密匙对来对启动器进行身份验证。ESX/ESXi 支持 CHAP 协议，VMware 建议将此协议用于 SAN 实施。要使用 CHAP 身份验证，ESX/ESXi 主机和 iSCSI 存储系统必须启用 CHAP，并具有公共凭据。

访问控制

访问控制是在 iSCSI 存储系统中设置的策略。大多数实施都支持以下三种访问控制类型的一种或多种：

- 通过启动器名称
- 通过 IP 地址
- 通过 CHAP 协议

仅符合所有规则的启动器才可以访问 iSCSI 卷。

错误纠正

为保护 iSCSI 标头和数据的完整性，iSCSI 协议定义了错误纠正方法，这些方法称为标头摘要和数据摘要。

这两类参数默认情况下为禁用状态，但可以启用它们。这些摘要分别适用于 iSCSI 启动器和目标之间双向传输的标头和 SCSI 数据。

除其他网络层（例如 TCP 和以太网）提供的完整性检查外，标头和数据摘要还检查端到端、非加密数据的完整性。它们检查完整的通信路径，包括可更改网络级流量的所有因素，例如路由器、交换机和代理。

建立 iSCSI 连接之后，将协商摘要的存在和类型。启动器和目标就摘要配置达成一致后，此摘要必须用于它们之间的所有流量。

启用标头和数据摘要要求对启动器和目标进行额外处理，并可影响吞吐量和 CPU 使用性能。

注意 使用 Intel Nehalem 处理器的系统会卸载 iSCSI 摘要计算功能，从而减少对性能的影响。

将 ESX/ESXi 与 SAN 配合使用概述

将 ESX/ESXi 与 SAN 配合使用可提高灵活性、效率以及可靠性。将 ESX/ESXi 与 SAN 配合使用还支持集中式管理、故障切换和负载平衡技术。

下面是将 ESX/ESXi 与 SAN 配合使用的优点：

- 您可安全地存储数据并配置多个指向存储器的路径，从而避免出现单一故障点。
- 将 SAN 与 ESX/ESXi 系统配合使用可将故障恢复功能扩展到服务器。使用 SAN 存储器时，所有应用程序都可在原始主机出现故障后在其他主机上立即重新启动。
- 可以使用 VMware vMotion 执行虚拟机的实时迁移。
- 如果虚拟机的主机出现故障，可将 VMware High Availability (HA) 与 SAN 结合使用，在另一台服务器上以虚拟机的最后已知状态重新启动虚拟机。
- 可使用 VMware 容错 (FT) 复制两台不同主机上的受保护虚拟机。如果首选主机出现故障，虚拟机可在辅助主机上继续不间断地运行。
- 使用 VMware Distributed Resource Scheduler (DRS) 将虚拟机从一台主机迁移到另一台主机以实现负载平衡。由于存储器位于共享 SAN 阵列，因此应用程序可继续无缝运行。
- 如果使用 VMware DRS 群集，请将 ESX/ESXi 主机置于维护模式，以便系统将所有正在运行的虚拟机迁移到其他 ESX/ESXi 主机。随后便可在原始主机上执行升级或其他维护操作。

VMware 虚拟机的可移植性和封装与此存储器的共享特性相辅相成。当虚拟机位于基于 SAN 的存储器上时，您可以快速关闭一台服务器上的虚拟机并在另一台服务器上将其启动，或在一台服务器上将其挂起并在同一网络上的另一台服务器上恢复操作。此功能可实现在迁移计算资源的同时保持一致的共享访问。

ESX/ESXi 和 SAN 用例

将 ESX/ESXi 与 SAN 配合使用时，可以执行多项任务。

将 ESX/ESXi 与 SAN 配合使用对于以下任务非常有效：

零停机时间维护
在执行 ESX/ESXi 主机或基础结构维护时，使用 VMware DRS 或 vMotion 将虚拟机迁移至其他服务器。如果共享存储器位于 SAN 上，您无需中断虚拟机的用户服务便可执行维护。

负载平衡
使用 vMotion 或 VMware DRS 将虚拟机迁移至其他主机以实现负载平衡。如果共享存储器位于 SAN 上，您无需中断虚拟机的用户服务便可执行负载平衡。

存储器整合与存储器布局的简化
如果使用多台主机，且每台主机都运行多个虚拟机，那么当主机上的存储器不再够用时将需要使用外部存储器。选择 SAN 作为外部存储器可使系统架构更为简单，同时还有其他优点。

先预留一个较大的卷，然后根据需要将各部分分配给虚拟机。从存储设备分配及创建卷需要一次完成。

灾难恢复
将所有数据存储在 SAN 上为数据备份的远程存储提供了便利。如果某一站点受到损坏，可以在远程 ESX/ESXi 主机上重新启动虚拟机进行恢复。

简化阵列迁移和存储升级
在购买新存储系统或阵列时，使用存储 vMotion 执行实时自动迁移，将虚拟机磁盘文件从现有存储器迁移到新的目标，且不中断虚拟机的用户。

查看详细信息

除本文档外，另有许多其他资源可帮助您配置 ESX/ESXi 系统与 SAN 配合使用：

- 对于大多数设置问题，请使用存储阵列供应商的文档。存储阵列供应商可能还会提供介绍在 ESX/ESXi 环境中使用存储阵列的文档。
- VMware 文档网站。
- 《光纤通道 SAN 配置指南》论述了 ESX/ESXi 与光纤通道存储区域网络配合使用的情况。
- 《VMware I/O 兼容性指南》列出了当前已认可的 HBA、HBA 驱动程序以及驱动程序版本。
- 《VMware 存储器/SAN 兼容性指南》列出了当前已认可的存储阵列。
- 《VMware 发行说明》提供了有关已知问题和解决办法的信息。
- 《VMware 知识库》包含有关常见问题和解决办法的信息。

将 SAN 存储器与 ESX/ESXi 配合使用的细节

将 SAN 与 ESX/ESXi 主机一起使用在很多方面有别于传统的 SAN 使用情况。

将 SAN 存储器与 ESX/ESXi 配合使用时，请记住以下注意事项：

- 您无法直接访问使用该存储器的虚拟机操作系统。使用传统工具，只能监控 VMware ESX/ESXi 操作系统。可以使用 vSphere Client 监控虚拟机。
- 通过 SAN 管理工具可看到的 HBA 属于 ESX/ESXi 系统（而不是属于虚拟机）。
- ESX/ESXi 系统可执行多路径。

第三方管理应用程序

可以将第三方管理应用程序与 ESX/ESXi 主机配合使用。

大多数 iSCSI 存储器硬件都附送存储管理软件。在许多情况下，此软件是一个 Web 应用程序，可与连接到网络的任何 Web 浏览器配合使用。在其他情况下，此软件通常在存储系统或单个服务器上运行，独立于使用 SAN 作为存储器的服务器。

此第三方管理软件可用于执行下列任务：

- 存储阵列管理（包括 LUN 创建、阵列缓存管理、LUN 映射以及 LUN 安全）。
- 设置复制、检查点、快照或镜像。

如果决定在虚拟机上运行 SAN 管理软件，您可享有运行虚拟机的一系列优点，包括使用 vMotion 和 VMware HA 进行故障切换。但是，由于附加了间接级别，管理软件可能看不到 SAN。在这种情况下，可以使用 RDM。

注意 虚拟机能否成功运行管理软件取决于特定的存储系统。

了解 VMFS 数据存储

ESX/ESXi 使用数据存储来存储虚拟磁盘，数据存储是一种逻辑容器，它将存储器的细节针对虚拟机隐藏起来，并提供一个统一的模型来存储虚拟机文件。在存储设备上部署的数据存储通常使用 VMware 虚拟机文件系统 (VMFS) 格式，该格式是针对存储虚拟机而优化的特殊高性能文件系统格式。

VMFS 数据存储可运行多台虚拟机。VMFS 为虚拟机文件提供了分布式锁定功能，因此，在多台 ESX/ESXi 主机共享同一 VMFS 数据存储的 SAN 环境中，虚拟机也能安全地工作。

使用 vSphere Client，可以在 ESX/ESXi 主机发现的基于块的存储设备上预先设置 VMFS 数据存储。一个 VMFS 数据存储可扩展为多个物理存储数据区，其中包括 SAN LUN 和本地存储器。使用此功能可以将存储放在存储池中，并灵活地创建虚拟机所需的数据存储。

当虚拟机在数据存储上运行时，您可以增加数据存储的容量。这一功能可让您在虚拟机需要新空间时为 VMFS 数据存储添加新空间。VMFS 专用于从多台物理机进行的并发访问，并在虚拟机文件上执行适当的访问控制。

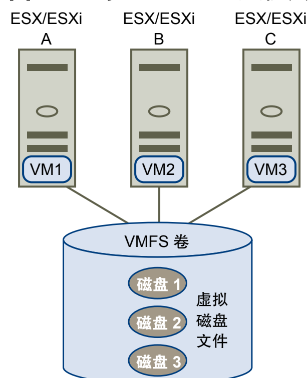
跨 ESX/ESXi 主机共享 VMFS 数据存储

作为一个群集文件系统，VMFS 允许多个 ESX/ESXi 主机同时访问同一个 VMFS 数据存储。

为了确保多台服务器不会同时访问同一个虚拟机，VMFS 提供了磁盘锁定。

图 1-3 显示了共享同一 VMFS 卷的若干 ESX/ESXi 系统。

图 1–3 跨 ESX/ESXi 主机共享 VMFS 数据存储



由于虚拟机共享一个公共的 VMFS 数据存储，因此可能很难得出固定的高峰访问时段或优化性能。必须针对高峰时段规划虚拟机存储访问，但是不同的应用程序可能具有不同的高峰访问时段。VMware 建议您使服务器、CPU 及存储器上的虚拟机达到负载平衡。在每台服务器上都运行多台虚拟机的组合，以便同一区域不会同时面临很高的资源需求。

元数据更新

VMFS 数据存储可保存虚拟机文件、目录、符号链接、RDM 描述符文件等等。数据存储还可保存这些对象的所有映射信息的一致性视图。此映射信息称为元数据。

元数据会在每次访问虚拟机文件的属性时更新或在执行以下任一操作时修改：

- 创建、扩展或锁定虚拟机文件
- 更改文件的属性
- 启动或关闭虚拟机

决定 LUN 的大小和数目

在使用 VMFS 数据存储格式化 LUN 之前，必须规划如何设置 ESX/ESXi 系统的存储器。

决定 LUN 的大小和数目时，请牢记以下注意事项：

- 每个 LUN 应具有正确的 RAID 级别和存储特性，适用于在使用该 LUN 的虚拟机中运行的应用程序。
- 一个 LUN 必须仅包含一个 VMFS 数据存储。
- 如果多台虚拟机访问同一个 VMFS，请使用磁盘份额区分虚拟机的优先级。

出于以下原因，您可能需要较少、较大的 LUN：

- 在不向存储管理员要求更多空间的情况下，使创建虚拟机的灵活性更大。
- 使调整虚拟磁盘大小、执行快照等等的灵活性更大。
- 使要管理的 VMFS 数据存储变得更少。

出于以下原因，您可能需要较多、较小的 LUN：

- 减少存储空间浪费。
- 不同的应用程序可能需要不同的 RAID 特性。
- 为每个 LUN 设置多路径策略和磁盘份额，从而增加灵活性。

- 使用 Microsoft 群集服务要求每个群集磁盘资源位于各自的 LUN 中。
- 由于对单个卷的争用情况较少，因此性能更佳。

当虚拟机的存储特性不可用时，通常没有简单的方法来确定要置备的 LUN 的数目和大小。可以使用预测性或自适应性方案决定 LUN 大小和数量。

使用预测性方案来决定 LUN 的大小和数目

设置 ESX/ESXi 系统的存储器时，必须在创建 VMFS 数据存储之前决定要置备的 LUN 的大小和数目。可以使用预测性方案决定 LUN 的大小和数目。

步骤

1. 置备若干具有不同存储特性的 LUN。
2. 在每个 LUN 上创建一个 VMFS 数据存储，并依照相应特性标记各个数据存储。
3. 创建虚拟磁盘以包含在 LUN 上创建的 VMFS 数据存储中的虚拟机应用程序数据，这些 LUN 具有满足其上应用程序要求的 RAID 级别。
4. 使用磁盘份额来区别高优先级虚拟机与低优先级虚拟机。

   注意　磁盘份额仅表示给定主机内的比例。分配给某一主机上虚拟机的份额并不影响其他主机上的虚拟机。

5. 运行应用程序以确定虚拟机性能是否可接受。

使用自适应性方案来决定 LUN 的大小和数目

设置 ESX/ESXi 主机的存储器时，必须在创建 VMFS 数据存储之前决定要置备的 LUN 的数目和大小。可以使用自适应性方案决定 LUN 的大小和数目。

步骤

1. 置备一个较大的 LUN（RAID 1+0 或 RAID 5），同时启用写入缓存。
2. 在该 LUN 上创建 VMFS。
3. 在 VMFS 上创建四个或五个虚拟磁盘。
4. 运行应用程序以确定磁盘性能是否可接受。

如果性能可接受，您可以在 VMFS 上再添加虚拟磁盘。如果性能不可接受，请新建大的 LUN（可能采用其他 RAID 级别）并重复此过程。使用迁移，以防止在重新创建 LUN 时丢失虚拟机数据。

使用磁盘份额区分虚拟机优先级

如果多台虚拟机访问同一 VMFS 数据存储（因此也是访问同一 LUN），请使用磁盘份额区分虚拟机对磁盘的访问优先级。磁盘份额可区别高优先级虚拟机与低优先级虚拟机。

步骤

1. 启动 vSphere Client 并连接到 vCenter Server。
2. 在“清单”面板中选择虚拟机，并从菜单中选择**编辑设置**。
3. 单击**资源**选项卡并单击**磁盘**。

4 双击要修改的磁盘的**份额**列，然后在下拉菜单中选择所需值。

份额是表示用于控制所有虚拟机磁盘带宽的相对衡量指标的值。将“低”、“正常”、“高”及“自定义”值与服务器、ESX 主机和服务控制台中所有虚拟机的所有份额的总数进行比较。份额分配符号值可用于将其转换配置为数字值。

5 单击**确定**保存选择。

注意 磁盘份额仅表示给定 ESX/ESXi 主机内的比例。分配给某一主机上虚拟机的份额并不影响其他主机上的虚拟机。

虚拟机如何访问 SAN 上的数据

ESX/ESXi 会将虚拟机的磁盘文件存储到位于 SAN 存储设备上的 VMFS 数据存储内。当虚拟机客户机操作系统向其虚拟磁盘发出 SCSI 命令时，SCSI 虚拟化层会将这些命令转换成 VMFS 文件操作。

虚拟机与其在 SAN 上存储的虚拟磁盘交互时，将发生以下过程：

1 虚拟机中的客户机操作系统读写 SCSI 磁盘时，将向虚拟磁盘发出 SCSI 命令。

2 虚拟机操作系统中的设备驱动程序将与虚拟 SCSI 控制器进行通信。

3 虚拟 SCSI 控制器将命令转发至 VMkernel。

4 VMkernel 将执行以下任务。

- 在 VMFS 卷中查找与客户机虚拟机磁盘对应的文件。
- 将对虚拟磁盘上块的请求映射到相应物理设备上的块。
- 将修改后的 I/O 请求从 VMkernel 中的设备驱动程序发送至 iSCSI 启动器（硬件或软件）。

5 如果 iSCSI 启动器是硬件 iSCSI 适配器（独立或从属），则该适配器执行以下任务：

- 将 I/O 请求封装到 iSCSI 协议数据单元 (Protocol Data Unit, PDU)。
- 将 iSCSI PDU 封装到 TCP/IP 数据包。
- 通过以太网将 IP 数据包发送至 iSCSI 存储系统。

6 如果 iSCSI 启动器是软件 iSCSI 适配器，则将发生以下事件。

- iSCSI 启动器将 I/O 请求封装到 iSCSI PDU 中。
- 启动器通过 TCP/IP 连接发送 iSCSI PDU。
- VMkernel TCP/IP 堆栈将 TCP/IP 数据包转发到物理网卡。
- 物理网卡通过以太网将 IP 数据包发送至 iSCSI 存储系统。

7 根据 iSCSI 启动器用于连接网络的端口，以太网交换机和路由器将请求递送至主机要访问的存储设备。

了解多路径和故障切换

要维持 ESX/ESXi 主机和其存储器之间的持续连接，ESX/ESXi 必须支持多路径。通过多路径技术，用户可以使用多个物理路径在主机和外部存储设备之间传输数据。

如果 SAN 网络中的任一网元（如适配器、交换机或线缆）发生故障，则 ESX/ESXi 可以切换到另一不使用故障组件的物理路径。这种避免使用故障组件的路径切换过程称为路径故障切换。

除路径故障切换外，多路径还提供负载平衡。负载平衡是在多个物理路径中分配 I/O 负载的过程。负载平衡可以减少或消除潜在的瓶颈。

注意 在路径故障切换发生时，虚拟机 I/O 最多能延迟 60 秒。此延迟使 SAN 可以在拓扑更改后稳定其配置。通常，主动-被动阵列上的 I/O 延迟时间可能会更长，而在主动-主动阵列上则更短。

管理多路径

为管理存储多路径，ESX/ESXi 使用特殊的 VMkernel 层（即，可插入存储架构 (PSA)）。PSA 是一个开放式模块框架，用于协调多个多路径插件 (MPP) 的同时操作。

ESX/ESXi 默认情况下提供的 VMkernel 多路径插件是 VMware 本机多路径插件 (NMP)。NMP 是管理子插件的可扩展模块。NMP 子插件有两种类型，即存储阵列类型插件 (SATP) 和路径选择插件 (PSP)。SATP 和 PSP 可以是 VMware 提供的内置插件，也可以由第三方提供。

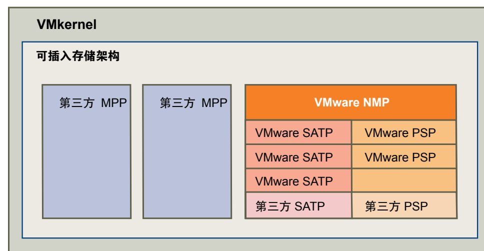
如果需要更多多路径功能，第三方还可以提供 MPP 以作为默认 NMP 的附属或替代运行。

当协调 VMware NMP 和所安装的任何第三方 MPP 时，PSA 将执行以下任务：

- 加载和卸载多路径插件。
- 对特定插件隐藏虚拟机细节。
- 将特定逻辑设备的 I/O 请求路由到管理该设备的 MPP。
- 处理逻辑设备的 I/O 排队操作。
- 在虚拟机之间实现逻辑设备带宽共享。
- 处理物理存储 HBA 的 I/O 排队操作。
- 处理物理路径发现和移除。
- 提供逻辑设备和物理路径 I/O 统计信息。

如图 1-4 所示，多个第三方 MPP 可以与 VMware NMP 并行运行。安装后，第三方 MPP 将替代 NMP 的行为，并且完全控制指定存储设备的路径故障切换和负载平衡操作。

图 1–4 可插入存储架构



多路径模块执行以下操作：

- 管理物理路径声明和取消声明。
- 管理逻辑设备的创建、注册和取消注册。
- 将物理路径与逻辑设备关联。
- 支持路径故障检测和修复。
- 处理逻辑设备的 I/O 请求：
  - 为请求选择最佳物理路径。
  - 根据存储设备，执行处理路径故障和 I/O 命令重试所需的特定操作。
- 支持管理任务，如中止或重置逻辑设备。

VMware 多路径模块

默认情况下，ESX/ESXi 提供名为本机多路径插件 (NMP) 的可扩展多路径模块。

一般来说，VMware NMP 支持 VMware 存储 HCL 上列出的所有存储阵列，并基于阵列类型提供默认的路径选择算法。还将一组物理路径与特定存储设备或 LUN 关联。存储阵列类型插件 (SATP) 负责处理给定存储阵列的路径故障切换。路径选择插件 (PSP) 负责确定哪个物理路径用于向存储设备发出 I/O 请求。SATP 和 PSP 是 NMP 模块中的子插件。

安装 ESX/ESXi 时，会自动安装适合所用阵列的 SATP。您不需要获取或下载任何 SATP。

VMware SATP

存储阵列类型插件 (SATP) 与 VMware NMP 一起运行，负责特定于阵列的操作。

ESX/ESXi 为 VMware 支持的各种类型阵列提供 SATP。它还提供默认 SATP 以支持非特定的主动-主动和 ALUA 存储阵列，以及适用于直接连接的设备的本地 SATP。每个 SATP 适合特定类别的存储阵列的特殊特性，并可以执行检测路径状况和激活被动路径所需的特定于阵列的操作。因此，NMP 模块本身可以使用多个存储阵列，而无需了解存储设备的特性。

在 NMP 确定要为特定存储设备使用哪个 SATP 并将该 SATP 与存储设备的物理路径相关联之后，该 SATP 会执行以下任务：

- 监控每个物理路径的健康状况。
- 报告每个物理路径的状况变化。
- 执行存储器故障切换所需的特定于阵列的操作。例如，对于主动-被动设备，它可以激活被动路径。

VMware PSP

路径选择插件 (PSP) 与 VMware NMP 一起运行，负责选择 I/O 请求的物理路径。

VMware NMP 根据与每个逻辑设备物理路径关联的 SATP 为该设备分配默认 PSP。可以替代默认 PSP。

默认情况下，VMware NMP 支持以下 PSP：

最近使用 (VMW_PSP_MRU)
选择 ESX/ESXi 主机最近用于访问指定设备的路径。如果此路径不可用，则主机会切换到替代路径并在该新路径可用时继续使用它。MRU 是主动-被动阵列的默认路径策略。

固定 (VMW_PSP_FIXED)
使用指定首选路径（如果已配置）。否则，它将使用在系统引导时间发现的第一个工作路径。如果主机不能使用首选路径，则它会选择随机替代可用路径。一旦首选路径可用，主机便会恢复到首选路径。“固定”是主动-主动阵列的默认路径策略。



小心 如果用于主动-被动阵列，则**固定**路径策略可能会导致路径抖动。

VMW_PSP_FIXED_AP
将“固定”功能扩展到主动-被动阵列和 ALUA 模式阵列。

循环 (VMW_PSP_RR)
使用路径选择算法轮流选择所有可用的活动路径，并在路径之间启用负载平衡。

VMware NMP I/O 流

当虚拟机向 NMP 管理的存储设备发出 I/O 请求时，将发生以下过程。

1 NMP 调用分配给此存储设备的 PSP。
2 PSP 将选择要通过其发出 I/O 的相应物理路径。
3 NMP 在 PSP 选择的路径上发出 I/O 请求。

4 如果 I/O 操作成功，则 NMP 报告其完成。

5 如果 I/O 操作报告错误，则 NMP 调用适当的 SATP。

6 SATP 解释 I/O 命令错误，并在适当时激活非活动路径。

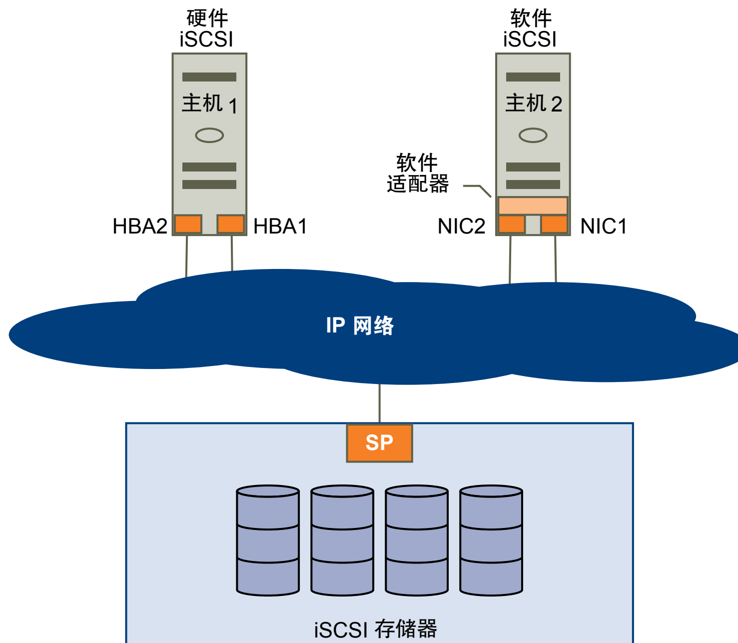
7 此时将调用 PSP 以选择要通过其发出 I/O 的新路径。

基于主机的路径故障切换

在设置 ESX/ESXi 主机以用于多路径和故障切换时，可以根据主机上 iSCSI 启动器的类型使用多个 iSCSI HBA 或多个网卡。

图 1–5 显示了对于不同类型的 iSCSI 启动器可能的多路径设置。

图 1–5 基于主机的路径故障切换



硬件 iSCSI 故障切换

在硬件 iSCSI 中，主机通常具有两个或更多的可用硬件 iSCSI 适配器，使用一个或多个交换机可从这些适配器到达存储系统。或者，设置可能包括一个适配器和两个存储处理器，以便适配器可以使用不同的路径到达存储系统。

如图 1–5 所示，主机 1 有两个硬件 iSCSI 适配器（HBA1 和 HBA2），用于提供两条到存储系统的物理路径。主机上的多路径插件，不论是 VMkernel NMP 还是任何第三方 MPP，默认情况下都可以访问这些路径，并可以监控每个物理路径的健康状况。例如，如果 HBA1 或 HBA1 与网络之间的链路发生故障，多路径插件可以将路径切换到 HBA2。

软件 iSCSI 故障切换

如图 1–5 的主机 2 所示，通过软件 iSCSI，您可以使用多张网卡为主机和存储系统间的 iSCSI 连接提供故障切换和负载平衡功能。

对于此设置，因为多路径插件对于主机上的物理网卡没有直接访问权，所以您首先需要将每张物理网卡连接至单独的 VMkernel 端口。然后使用端口绑定技术将所有的 VMkernel 端口与软件 iSCSI 启动器相关联。因此，连接到单独网卡的每个 VMkernel 端口将成为 iSCSI 存储堆栈及其存储感知多路径插件可使用的另一条路径。

有关如何为软件 iSCSI 配置多路径的详细信息，请参见第 31 页，“软件 iSCSI 和从属硬件 iSCSI 的网络配置”。

基于阵列的故障切换

一些 iSCSI 存储系统可自动管理其端口的路径使用情况，且这种管理路径使用情况的方式对 ESX/ESXi 是透明的。

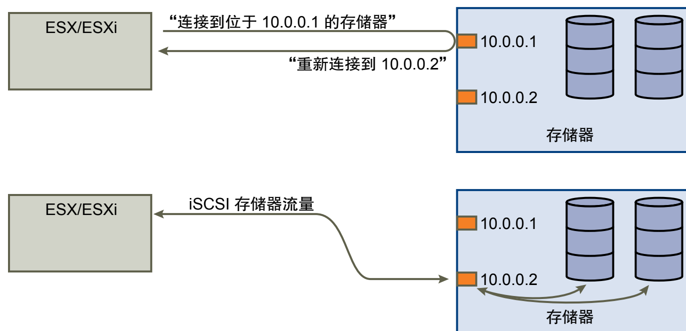
使用其中一个存储系统时，ESX/ESXi 不会看到存储器上的多个端口，并且不能选择其连接的存储器端口。这些系统只有一个可供 ESX/ESXi 进行初始通信的虚拟端口地址。在初始通信期间，存储系统可重定向 ESX/ESXi，使其与存储系统上的其他端口进行通信。ESX/ESXi 中的 iSCSI 启动器遵守此重新连接请求，并与系统上的其他端口连接。存储系统使用此技术在可用端口间分散负载。

如果 ESX/ESXi 丢失其中一个端口的连接，则会自动尝试与存储系统的虚拟端口重新连接，并应重定向至活动的可用端口。此重新连接和重定向将迅速发生，一般不会中断正在运行的虚拟机。这些存储系统还可请求 iSCSI 启动器重新连接系统，以更改其连接的存储器端口。这样可最高效地使用多个端口。

图 1-6 显示端口重定向的示例。ESX/ESXi 尝试连接 10.0.0.1 虚拟端口。此存储系统将此请求重定向至 10.0.0.2。ESX/ESXi 与 10.0.0.2 连接并使用此端口进行 I/O 通信。

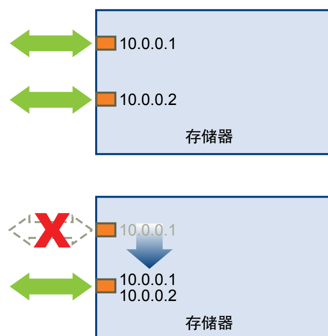
注意 此存储系统并不总是重定向连接。端口 10.0.0.1 也可用于流量。

图 1–6 端口重定向



如果存储系统上充当虚拟端口的端口变得不可用，则存储系统会将虚拟端口地址重新分配至系统上的其他端口。图 1-7 显示了此类端口重新分配的示例。在此例中，虚拟端口 10.0.0.1 变得不可用，存储系统将虚拟端口 IP 地址重新分配至其他端口。第二个端口对两个地址做出响应。

图 1–7 端口重新分配



对于基于阵列的故障切换，只有在使用 ESX/ESXi 主机上的多个端口时，才可以有多条到存储器的路径。这些路径是主动-主动路径。有关详细信息，请参见第 42 页，“iSCSI 会话管理”。

选择虚拟机位置

如果您要设法优化虚拟机的性能，存储位置是个重要因素。提供高性能和高可用性但价格昂贵的存储器与性能较低但成本也较低的存储器之间始终存在一个权衡。

根据多种因素，存储器可分为不同级别：

- 高端。提供高性能和高可用性。可能提供内置快照，便于备份及时间点 (PiT) 还原。支持复制、完全 SP 冗余和 SAS 驱动器。使用高成本心轴。
- 中端。提供中等程度的性能、较低可用性、部分 SP 冗余以及 SCSI 或 SAS 驱动器。可能提供快照。使用中等成本的心轴。
- 低端。提供低性能及少许内部存储冗余。使用低端 SCSI 驱动器或 SATA（串行低成本心轴）。

并非所有应用程序都需要性能最高、可用性最佳的存储器，至少不是在整个生命周期内一直需要。

注意 如果需要使用某些高端功能（如快照），但不希望额外支出费用，则可以在软件中实现某些高性能特性。例如，您可以在软件中创建快照。

决定放置某虚拟机的位置时，请考虑以下问题：

- 该虚拟机的重要程度如何？
- 有哪些性能及可用性要求？
- 有哪些 PiT 还原要求？
- 有哪些备份要求？
- 有哪些复制要求？

由于重要程度发生改变或技术发展使得目前的较高端功能降为低端功能，虚拟机在其整个生命周期过程中可能变更级别。重要程度是相对的，并且可能由于多种原因（包括组织、操作流程、法规要求及灾难规划等方面的变化）而改变。

针对服务器故障的设计

SAN 存储器的 RAID 架构本质上是在物理磁盘级别提供故障保护。SAN 在服务器与存储器之间提供多个路径，可防止网络或端口故障。使整个环境变得耐故障的最后一步即提供针对服务器故障的保护。

使用 VMware HA

ESX/ESXi 提供的其中一个故障切换选项是 VMware High Availability (HA)。

使用 VMware HA 可将虚拟机组织为故障切换组。某一主机发生故障时，其所有虚拟机将立即在其他主机上启动。虚拟机在其他主机上还原后会丢失内存状况，但磁盘状况与主机发生故障时完全一致（崩溃一致故障切换）。HA 需要共享存储器（如 SAN）。

注意 您必须获得许可才能使用 VMware HA。

服务器故障切换和存储器注意事项

如果要将 ESX/ESXi 主机配置为与 SAN 协同工作，必须使整个环境变得耐故障并针对主机故障进行保护。

对于各种类型的服务器故障切换，您必须遵循下列做法：

- 仅当每台服务器都对同一存储器具有访问权限时用于服务器故障切换的方法才可行。因为多台服务器需要大量磁盘空间，而且存储系统故障切换与服务器故障切换相辅相成，所以通常会将 SAN 与服务器故障切换结合使用。
- 规划 SAN 与服务器故障切换联合使用时，群集虚拟机使用的所有数据存储均必须对所有 ESX/ESXi 主机可见。

  虽然主机可以访问数据存储，但并非该主机上的所有虚拟机就一定具有对数据存储上所有数据的访问权限。虚拟机只能访问为其配置的虚拟磁盘。假设出现配置错误，在虚拟机引导时，虚拟磁盘将被锁定，因此不会发生损坏现象。

注意 一般来说，从 SAN 引导时，应该只有从该卷引导的主机可以看到相应的引导卷。一种例外情况就是当您通过将另一主机指向同一个卷来尝试从故障中恢复时。在这种情况下，上述 SAN 卷并非真正从 SAN 中引导。因 SAN 已损坏，没有任何主机从其引导。SAN 卷是对主机可见的常规非引导卷。

LUN 的显示与重新扫描

SAN 是动态的，某特定主机可使用哪些 LUN 随多种因素而变化。

VMkernel 在启动时会发现 LUN，这些 LUN 随后即在 vSphere Client 中可见。如果对 LUN 进行了更改，您必须重新扫描才能看到相应更改。

- iSCSI 存储器上创建的新 LUN
- 对 LUN 访问控制的更改
- 连接性方面的更改

配置 iSCSI 启动器和存储器 2

必须先安装 iSCSI 启动器和存储器，ESX/ESXi 才能与 SAN 配合使用。

为此，必须首先遵循某些基本要求，然后按照安装和设置硬件或软件 iSCSI 启动器的最佳做法来访问 SAN。

本章讨论了以下主题：

ESX/ESXi iSCSI SAN 要求

必须满足多项要求，才能使 ESX/ESXi 主机与 SAN 一起正常工作。

- 验证是否支持 SAN 存储硬件和固件组合与 ESX/ESXi 系统配合使用。有关最新列表，请参见在线的《硬件兼容性指南》中的“存储器/SAN”部分。
- 将系统配置为每个 LUN 只有一个 VMFS 数据存储。使用 VMFS-3 时，无需设置可访问性。
- 除非使用无磁盘服务器（从 SAN 引导），否则请勿在 SAN LUN 上设置诊断分区。共享的诊断分区适用于从 SAN 引导无磁盘服务器的情况。

- 使用 RDM 访问任何裸磁盘。
- 设置客户机操作系统中的 SCSI 控制器驱动程序以指定一个足够大的队列。您可以在系统设置过程中设置物理 HBA 的队列深度。
- 在运行 Microsoft Windows 的虚拟机上，增加 SCSI `TimeoutValue` 参数值，以使 Windows 能接受因路径故障切换所导致的更长 I/O 延迟。

ESX/ESXi iSCSI SAN 限制

本主题列出了将 ESX/ESXi 与 iSCSI SAN 配合使用时存在的限制。

- ESX/ESXi 不支持 iSCSI 连接磁带设备。
- 您不能使用虚拟机多路径软件对单个物理 LUN 执行 I/O 负载平衡。
- 当主机同时使用从属和独立硬件 iSCSI 适配器时，它不能访问同一 LUN。

设置 LUN 分配

准备让 ESX/ESXi 系统使用 iSCSI SAN 存储器时，需要设置 LUN 分配。

请注意以下几点：

- 存储器置备。为了确保 ESX/ESXi 主机在启动时能够识别 LUN，请配置所有 iSCSI 存储目标，以便主机可以对它们进行访问和使用。此外，请对主机进行配置，使其可以发现所有可用的 iSCSI 目标。
- vMotion 和 VMware DRS。当您使用 vCenter Server 以及 vMotion 或 DRS 时，请确保用于虚拟机的 LUN 已置备到所有的 ESX/ESXi 主机。此配置可为移动虚拟机提供最大的自由。
- 主动-主动与主动-被动阵列。将 vMotion 或 DRS 用于主动-被动 SAN 存储设备时，请确保所有 ESX/ESXi 系统通向所有存储处理器的路径均一致。否则在进行 vMotion 迁移时可能会导致路径抖动。

  对于在线的《VMware 兼容性指南》中的“存储器/SAN”部分中未列出的主动-被动存储阵列，VMware 不支持存储器-端口故障切换。必须将服务器连接到存储系统的主动端口。此配置可确保向 ESX/ESXi 主机呈现 LUN。

网络配置和身份验证

在 ESX/ESXi 能够发现 iSCSI 存储器之前，必须配置 iSCSI 启动器，并且还可能必须设置身份验证。

- 对于软件 iSCSI 和从属硬件 iSCSI，必须针对 VMkernel 对网络进行配置。可以使用 `vmkping` 实用程序来验证网络配置。对于硬件 iSCSI，必须为 HBA 配置网络参数（例如 IP 地址、子网掩码和默认网关）。
- 如有必要，请检查并更改默认启动器名称。
- 必须设置存储系统的发现地址，并应能使用 `vmkping` ping 此地址。
- 在启动器和存储系统端启用 CHAP 身份验证。启用身份验证后，其适用于尚未发现的所有目标，但不适用于已发现的目标。设置发现地址后，发现的新目标将被公开，且在当时即可用。

有关如何使用 `vmkping` 命令的详细信息，请搜索 VMware 知识库。

设置独立硬件 iSCSI 适配器

独立硬件 iSCSI 适配器是可通过 TCP/IP 访问 iSCSI 存储器的专用第三方适配器。此 iSCSI 适配器负责 ESX/ESXi 系统的所有 iSCSI 以及网络处理和管理。

独立硬件 iSCSI 适配器的设置和配置过程包含以下步骤：

1 检查适配器是否需要获得许可。

  请参见供应商文档。

2 安装适配器。

  有关安装信息，请参见供应商文档。

3 验证适配器是否安装正确。

  请参见第 27 页，"查看独立硬件 iSCSI 适配器"。

4 配置发现地址。

  请参见第 37 页，"配置 iSCSI 启动器的发现地址"。

5 配置 CHAP 参数。

  请参见第 38 页，"配置 iSCSI 适配器的 CHAP 参数"。

要使主机能够访问 iSCSI 存储器，必须先安装硬件 iSCSI 适配器并配置发现地址和 CHAP 参数。

查看独立硬件 iSCSI 适配器

可以查看硬件 iSCSI 适配器以验证它是否已正确安装并准备好进行配置。

前提条件

安装硬件 iSCSI 适配器后，它会显示在可供配置的存储适配器列表上。您可查看其属性。

所需特权：**主机.配置.存储器分区配置**

步骤

1 登录到 vSphere Client，然后在"清单"面板中选择主机。

2 单击**配置**选项卡，然后在"硬件"面板中单击**存储适配器**。

  如果已安装，则硬件 iSCSI 启动器应显示在存储适配器列表上。

3 选择要查看的启动器。

  此时会显示启动器的默认详细信息，包括型号、iSCSI 名称、iSCSI 别名、IP 地址及目标和路径信息。

4 单击**属性**。

  此时将显示"iSCSI 启动器属性"对话框。**常规**选项卡显示了启动器的附加特性。

现在可配置硬件启动器或更改其默认特性。

更改独立硬件启动器的名称和 IP 地址

在配置独立硬件 iSCSI 启动器时，请确保其名称和 IP 地址的格式正确。

步骤

1 访问"iSCSI 启动器属性"对话框。

2 单击**配置**。

3 要更改启动器的默认 iSCSI 名称，请输入新的名称。

   确保所输入的名称在整个环境中唯一且其格式正确；否则，某些存储设备可能无法识别硬件 iSCSI 启动器。

4 （可选）输入 iSCSI 别名。

   别名是用于标识硬件 iSCSI 启动器的名称。

5 更改默认 IP 设置。

   必须更改默认 IP 设置，以便 IP SAN 的 IP 设置正确无误。与网络管理员一起确定 HBA 的 IP 设置。

6 单击**确定**保存更改。

如果更改 iSCSI 名称，该名称会在新的 iSCSI 会话中使用。但对于现有会话，直到注销并重新登录之后才能使用新设置。

设置和配置从属硬件 iSCSI 适配器

从属硬件 iSCSI 适配器是基于 VMware 网络以及由 VMware 提供的 iSCSI 配置和管理界面的第三方适配器。

这种适配器类型可以是一张为同一端口提供标准网络适配器和 iSCSI 卸载功能的卡，例如 Broadcom 5709 网卡。iSCSI 卸载功能作为 iSCSI 适配器显示在存储适配器列表中。虽然默认情况下 iSCSI 适配器已启用，但若要使其正常工作，必须设置适用于 iSCSI 流量的网络，并将适配器与相应的 VMkernel iSCSI 端口绑定。然后，可以配置此适配器。

从属硬件 iSCSI 适配器的完整设置和配置过程包括以下步骤：

1 查看从属硬件适配器。

   请参见第 29 页，“查看从属硬件 iSCSI 适配器”。

   如果从属硬件适配器未显示在存储适配器列表中，请检查是否需要获取它们的许可。请参见供应商文档。

2 确定从属硬件适配器与物理网卡之间的关联。

   请参见第 29 页，“确定从属硬件 iSCSI 与物理网络适配器之间的关联”

   确保记下对应物理网卡的名称。例如，vmhba33 适配器与 vmnic1 相对应，而 vmhba34 与 vmnic2 相对应。

3 通过创建用于 iSCSI 流量的端口来配置 iSCSI 网络。

   请参见第 31 页，“软件 iSCSI 和从属硬件 iSCSI 的网络配置”。

   对每个网卡打开一个端口。例如，为 vmnic1 网卡创建 vmk1 端口，并为 vmnic2 创建 vmk2 端口。

4 将 iSCSI 端口绑定到对应的从属硬件 iSCSI 适配器。无论您有多个适配器还是只有一个适配器，都必须执行此步骤。

   请参见第 34 页，“将 iSCSI 端口绑定到 iSCSI 适配器”。

   在此示例中，将端口 vmk1 绑定到 vmhba33，并将端口 vmk2 绑定到 vmhba34。

5 配置发现地址。

   请参见第 37 页，“配置 iSCSI 启动器的发现地址”。

6 配置 CHAP 参数。

   请参见第 38 页，“配置 iSCSI 适配器的 CHAP 参数”。

从属硬件 iSCSI 注意事项

当将从属硬件 iSCSI 适配器与 ESX/ESXi 一起使用时，某些注意事项适用。

- 使用任何从属硬件 iSCSI 适配器时，与该适配器关联的网卡的性能报告可能显示很少活动或无活动，即使在 iSCSI 流量很大时也是如此。出现这种行为是因为 iSCSI 流量绕过常规网络堆栈。
- Broadcom iSCSI 适配器在硬件中执行数据重组，其缓冲区空间有限。在拥堵的网络中或在负载下使用 Broadcom iSCSI 适配器时，启用流量控制可避免性能下降。

  流量控制可管理两个节点之间的数据传输速率，以防止发送方的速度过快而超过接收方。为了获得最佳效果，请在 I/O 路径端点处、主机和 iSCSI 存储系统上启用流量控制。
- Broadcom iSCSI 适配器不支持 IPv6 和巨帧。

查看从属硬件 iSCSI 适配器

查看从属硬件 iSCSI 适配器以验证其加载正确。

如果从属硬件适配器未显示在存储适配器列表中，请检查是否需要获取它的许可。请参见供应商文档。

步骤

1. 登录到 vSphere Client，然后在“清单”面板中选择主机。
2. 单击**配置**选项卡，然后在“硬件”面板中单击**存储适配器**。

   如果已安装，则从属硬件 iSCSI 适配器应显示在存储适配器列表中。
3. 选择要查看的适配器，然后单击**属性**。

   “iSCSI 启动器属性”对话框会显示该适配器的默认详细信息，包括 iSCSI 名称和 iSCSI 别名。

确定从属硬件 iSCSI 与物理网络适配器之间的关联

需要确定与从属硬件 iSCSI 适配器关联的物理网卡的名称。还需要知道能够正确执行端口绑定的关联。

步骤

1. 使用 vSphere CLI 命令确定与 iSCSI 适配器关联的物理网卡的名称。

   ```
   esxcli swiscsi vmnic list -d vmhba#
   ```

   vmhba# 是 iSCSI 适配器的名称。
2. 在输出中，找到 `vmnic name:vmnic#` 这一行。

   vmnic# 是与 iSCSI 适配器对应的网络适配器的名称。

下一步

确定网卡的名称后，请在连接到网卡的 vSwitch 上创建一个 iSCSI 端口。然后将此端口绑定到从属硬件 iSCSI 适配器，以便您的主机可以指导通过网卡的 iSCSI 流量。

设置和配置软件 iSCSI 适配器

借助基于软件的 iSCSI 实施，可使用标准网卡将主机连接至 IP 网络上的远程 iSCSI 目标。ESX/ESXi 中内置的软件 iSCSI 适配器为此连接提供了便利，可通过网络堆栈与物理网卡进行通信。

使用 vSphere Client 连接 vCenter Server 或主机后，可以看到存储适配器列表上的软件 iSCSI 适配器。只显示一个软件 iSCSI 适配器。在使用软件 iSCSI 适配器之前，必须设置网络、启用适配器，并配置发现地址和 CHAP 等参数。软件 iSCSI 适配器配置工作流程包括以下步骤：

1 通过创建用于 iSCSI 流量的端口来配置 iSCSI 网络。

   请参见第 31 页，“软件 iSCSI 和从属硬件 iSCSI 的网络配置”。

2 启用软件 iSCSI 适配器。

   请参见第 30 页，“启用软件 iSCSI 适配器”。

3 如果将多个网卡用于软件 iSCSI 多路径，则通过将所有 iSCSI 端口连接到软件 iSCSI 适配器来执行端口绑定。

   请参见第 34 页，“将 iSCSI 端口绑定到 iSCSI 适配器”。

4 如有需要的话，请启用巨帧。必须通过 vSphere CLI 为每个 vSwitch 启用巨帧。

5 配置发现地址。

   请参见第 37 页，“配置 iSCSI 启动器的发现地址”。

6 配置 CHAP 参数。

   请参见第 38 页，“配置 iSCSI 适配器的 CHAP 参数”。

启用软件 iSCSI 适配器

必须启用软件 iSCSI 适配器，以便主机可以使用它来访问 iSCSI 存储器。

前提条件

在启用软件 iSCSI 适配器之前，必须设置 iSCSI 的网络。

注意 如果使用软件 iSCSI 适配器从 iSCSI 进行引导，则在首次引导时会启用该适配器，并自动创建网络配置。如果禁用了该适配器，它会在您每次引导主机时重新启用。

步骤

1 登录 vSphere Client，从“清单”面板中选择服务器。

2 单击**配置**选项卡，然后在“硬件”面板中单击**存储适配器**。

   此时将显示可用存储适配器的列表。

3 选择要配置的 iSCSI 启动器，然后单击**属性**。

4 单击**配置**。

5 要启用启动器，请选择**已启用**，然后单击**确定**。

启用启动器之后，主机将为其分配默认的 iSCSI 名称。如果需要您可以更改默认名称。

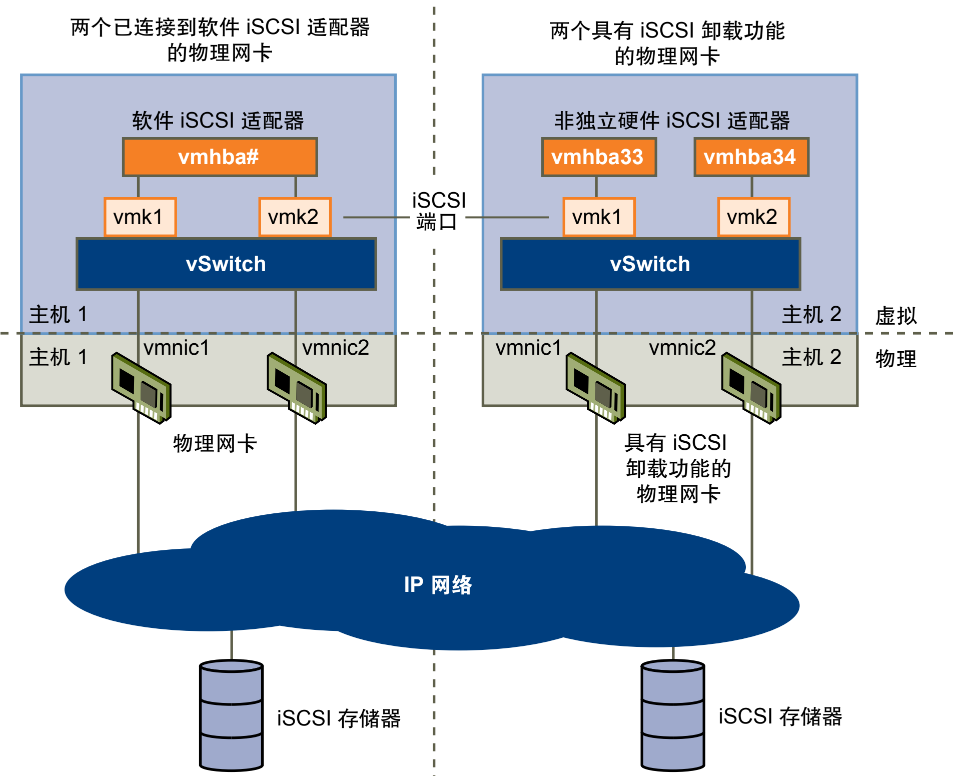
软件 iSCSI 和从属硬件 iSCSI 的网络配置

如果使用软件 iSCSI 适配器或从属硬件 iSCSI 适配器，必须先设置 iSCSI 网络，才能启用和配置 iSCSI 适配器。iSCSI 的网络配置涉及打开 VMkernel iSCSI 端口，用于 iSCSI 适配器与物理网卡之间的流量。

根据要用于 iSCSI 流量的物理网卡的数量，网络设置可能不同。

- 如果有单个物理网卡，请在与该网卡相连的 vSwitch 上创建一个 iSCSI 端口。VMware 建议为 iSCSI 指定一个单独的网络适配器。不要在 100 Mbps 或更慢的适配器上使用 iSCSI。
- 如果有两个或更多物理网卡用于 iSCSI，请为每个物理网卡创建一个单独的 iSCSI 端口，然后将这些网卡用于 iSCSI 多路径。请参见图 2-1。

图 2–1 iSCSI 网络



注意 使用从属硬件 iSCSI 适配器时，与该适配器关联的网卡的性能报告可能显示很少活动或无活动，即使在 iSCSI 流量很大时也是如此。出现这种行为是因为 iSCSI 流量绕过常规网络堆栈。

- 为单个网卡创建 iSCSI 端口第 32 页，

  使用此任务可将运行 iSCSI 存储器服务的 VMkernel 连接至物理网卡。如果只有一个物理网络适配器要用于 iSCSI 流量，则这是设置 iSCSI 网络所必须执行的唯一过程。
- 将多个网卡用于软件和从属硬件 iSCSI 第 32 页，

  如果主机有多个物理网卡用于 iSCSI，请使用 1:1 映射对每个物理网卡创建一个单独的 iSCSI 端口。
- 为多个网卡创建其他 iSCSI 端口第 33 页，

  如果您有两个或更多个网卡可指定给 iSCSI，而且希望将所有 iSCSI 网卡连接到单个 vSwitch，请使用此任务。在此任务中，需要使用 1:1 映射将 VMkernel iSCSI 端口与网络适配器相关联。

为单个网卡创建 iSCSI 端口

使用此任务可将运行 iSCSI 存储器服务的 VMkernel 连接至物理网卡。如果只有一个物理网络适配器要用于 iSCSI 流量，则这是设置 iSCSI 网络所必须执行的唯一过程。

步骤

1 登录 vSphere Client，在“清单”面板中选择主机。

2 依次单击**配置**选项卡和**网络**。

3 在“虚拟交换机”视图中，单击**添加网络**。

4 选择 **VMkernel**，然后单击**下一步**。

5 选择**创建虚拟交换机**以创建新的 vSphere。

6 选择要用于 iSCSI 流量的网卡。

重要事项 如果要为从属硬件 iSCSI 适配器创建一个端口，请确保选择与 iSCSI 组件相对应的网卡。请参见第 29 页，“确定从属硬件 iSCSI 与物理网络适配器之间的关联”。

7 单击**下一步**。

8 输入网络标签。

网络标签是用于识别所创建的 VMkernel 端口的友好名称，如 iSCSI。

9 单击**下一步**。

10 指定 IP 设置，然后单击**下一步**。

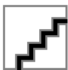
11 检查信息，然后单击**完成**。

将多个网卡用于软件和从属硬件 iSCSI

如果主机有多个物理网卡用于 iSCSI，请使用 1:1 映射对每个物理网卡创建一个单独的 iSCSI 端口。

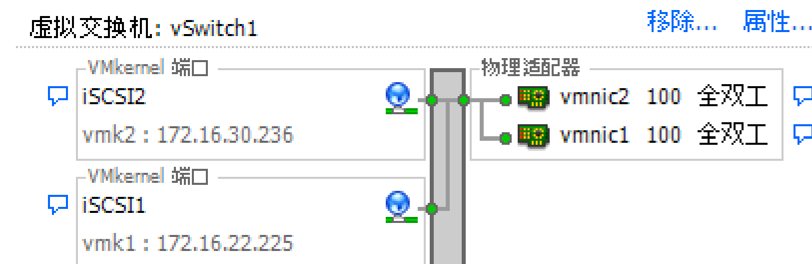
要实现 1:1 映射，请为每个网络适配器和 iSCSI 端口对指定一个单独的 vSwitch。请参见图 2-2。

图 2-2 单独的 vSwitch 上的 iSCSI 端口和网卡



替代方法是将所有网卡和 iSCSI 端口对添加到单个 vSwitch。请参见图 2-3。必须替代默认设置，并确保每个端口只映射到一个对应的活动网卡。

图 2-3 单个 vSwitch 上的 iSCSI 端口和网卡



有关将网卡和 VMkernel 端口对添加到 vSwitch 的详细信息，请参见第 33 页，“为多个网卡创建其他 iSCSI 端口”。

在将 iSCSI 端口映射到网络适配器之后，请使用 `esxcli` 命令将这些端口绑定到 iSCSI 适配器。对于从属硬件 iSCSI 适配器，无论您是使用一个网卡还是多个网卡，都要执行端口绑定。有关信息，请参见第 34 页，“将 iSCSI 端口绑定到 iSCSI 适配器”。

为多个网卡创建其他 iSCSI 端口

如果您有两个或更多个网卡可指定给 iSCSI，而且希望将所有 iSCSI 网卡连接到单个 vSwitch，请使用此任务。在此任务中，需要使用 1:1 映射将 VMkernel iSCSI 端口与网络适配器相关联。

现在需要将额外的网卡连接到现有 vSwitch，并将它们映射到相应的 iSCSI 端口。

注意 如果将具有多个 dvUplink 的 vNetwork 分布式交换机用于端口绑定，则需要为每个物理网卡创建一个单独的 dvPort 组。然后，应设置绑定策略，以便每个 dvPort 组只有一个活动 dvUplink。

有关 vNetwork 分布式交换机的详细信息，请参见《ESX/ESXi 配置指南》的“网络”部分。

前提条件

必须创建 vSwitch，将 iSCSI 端口映射到指定用于 iSCSI 流量的物理网卡。

步骤

1 登录 vSphere Client，在“清单”面板中选择主机。

2 依次单击**配置**选项卡和**网络**。

3 选择用于 iSCSI 的 vSwitch，然后单击**属性**。

4 将额外网络适配器连接到 vSwitch。

   a 在“vSwitch 属性”对话框中，单击**网络适配器**选项卡，然后单击**添加**。

   b 在列表中选择一个或多个网卡，然后单击**下一步**。

   对于从属硬件 iSCSI 适配器，确保只选择具有一个相应的 iSCSI 组件的那些网卡。

   c 检查“适配器摘要”页面上的信息，然后单击**完成**。

   此时将重新出现网络适配器列表，显示现在由 vSwitch 声明的网络适配器。

5 为已连接的所有网卡创建 iSCSI 端口。

   iSCSI 端口的数量必须与 vSwitch 上网卡的数量相对应。

   a 在“vSwitch 属性”对话框中，单击**端口**选项卡，然后单击**添加**。

   b 选择 **VMkernel**，然后单击**下一步**。

   c 在**端口组属性**下方，输入网络标签（如 iSCSI），然后单击**下一步**。

   d 指定 IP 设置，然后单击**下一步**。

   输入子网掩码时，请确保网卡已设置为其连接到的存储系统的子网。

   e 检查信息，然后单击**完成**。



小心 如果和 iSCSI 适配器（软件或从属硬件 iSCSI 适配器）配合使用的网卡与 iSCSI 目标不在同一子网中，则主机将无法建立从此网络适配器到目标的会话。

---

6 将每个 iSCSI 端口仅映射到一个活动网卡。

默认情况下，对于 vSwitch 上的每个 iSCSI 端口，所有网络适配器均显示为活动状态。必须替代此设置，以便每个端口只映射到一个相应的活动网卡。例如，iSCSI 端口 vmk1 映射到 vmnic1，端口 vmk2 映射到 vmnic2，等等。

a 在**端口**选项卡上，选择 iSCSI 端口，然后单击**编辑**。

b 单击**网卡绑定**选项卡，并选择**替代 vSwitch 故障切换顺序**。

c 仅将一个适配器指定为活动适配器，并将所有剩余的适配器移动到**未用的适配器**类别。

7 对 vSwitch 上的每个 iSCSI 端口重复最后一步。

下一步

执行此任务后，需要使用 `esxcli` 命令将 iSCSI 端口绑定到软件 iSCSI 或从属硬件 iSCSI 适配器。

将 iSCSI 端口绑定到 iSCSI 适配器

可将为网卡创建的 iSCSI 端口绑定到 iSCSI 适配器。对于软件 iSCSI 适配器，只有在您设置了两个或更多个用于 iSCSI 多路径的网卡时，才需要执行此任务。如果使用从属硬件 iSCSI 适配器，则无论您有多个适配器还是一个适配器，都需要执行此任务。

前提条件

您可完成以下任务：

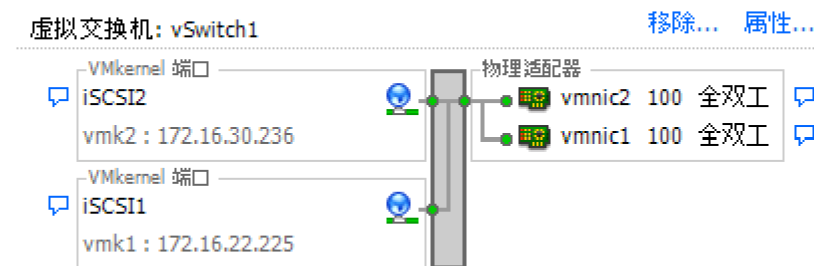
- 对于从属硬件 iSCSI 适配器，使物理网卡与 iSCSI 适配器之间正确地关联。请参见第 29 页，“查看从属硬件 iSCSI 适配器”。
- 设置用于 iSCSI 流量的网络。请参见第 31 页，“软件 iSCSI 和从属硬件 iSCSI 的网络配置”。
- 要使用软件 iSCSI 适配器，首先将它启用。请参见第 30 页，“启用软件 iSCSI 适配器”。

步骤

1 确定分配给物理网卡的 iSCSI 端口的名称。

vSphere Client 在网络标签下显示端口的名称。

在下图中，端口的名称为 vmk1 和 vmk2。



2 使用 vSphere CLI 命令将 iSCSI 端口绑定到 iSCSI 适配器。

esxcli swiscsi nic add -n *port_name* -d *vmhba*

重要事项 对于软件 iSCSI，对连接所有端口与软件 iSCSI 适配器的每个 iSCSI 端口重复此命令。对于从属硬件 iSCSI，确保将每个端口绑定到适当的相应适配器。

3 验证是否已将该端口添加到 iSCSI 适配器。

```
esxcli swiscsi nic list -d vmhba
```

4 使用 vSphere Client 重新扫描 iSCSI 适配器。

将 iSCSI 端口绑定到 iSCSI 适配器

查看有关示例，了解如何将为物理网卡创建的多个端口绑定到软件 iSCSI 适配器或多个从属硬件 iSCSI 适配器。

示例 2–1 将 iSCSI 端口连接到软件 iSCSI 适配器

此示例显示如何将 iSCSI 端口 vmk1 和 vmk2 连接到软件 iSCSI 适配器 vmhba33。

1 将 vmk1 连接到 vmhba33：**esxcli swiscsi nic add -n vmk1 -d vmhba33**。

2 将 vmk2 连接到 vmhba33：**esxcli swiscsi nic add -n vmk2 -d vmhba33**。

3 验证 vmhba33 配置：**esxcli swiscsi nic list -d vmhba33**。

即会列出 vmk1 和 vmk2。

如果通过 vSphere Client 显示 vmhba33 适配器的“路径”视图，则可以看见该适配器使用两个路径访问同一目标。路径的运行时名称是 vmhba33:C1:T1:L0 和 vmhba33:C2:T1:L0。此示例中的 C1 和 C2 表示用于多路径的两个网络适配器。

示例 2–2 将 iSCSI 端口连接到从属硬件 iSCSI 适配器

此示例显示如何将 iSCSI 端口 vmk1 和 vmk2 连接到对应的硬件 iSCSI 适配器 vmhba33 和 vmhba34。

1 将 vmk1 连接到 vmhba33：**esxcli swiscsi nic add -n vmk1 -d vmhba33**。

2 将 vmk2 连接到 vmhba34：**esxcli swiscsi nic add -n vmk2 -d vmhba34**。

3 验证 vmhba33 配置：**esxcli swiscsi nic list -d vmhba33**。

4 验证 vmhba34 配置：**esxcli swiscsi nic list -d vmhba34**。

断开 iSCSI 端口与 iSCSI 适配器的连接

如果需要更改用于 iSCSI 流量的网络配置（例如，移除网卡或 iSCSI 端口），应确保断开 iSCSI 端口与 iSCSI 适配器的连接。

重要事项 如果在主机与目标之间存在活动的 iSCSI 会话，则不能断开 iSCSI 端口的连接。

步骤

1 使用 vSphere CLI 断开 iSCSI 端口与 iSCSI 适配器的连接。

```
esxcli swiscsi nic remove -n port_name -d vmhba
```

2 验证是否已断开该端口与 iSCSI 适配器的连接。

```
esxcli swiscsi nic list -d vmhba
```

3 使用 vSphere Client 重新扫描 iSCSI 适配器。

将巨帧与 iSCSI 配合使用

巨帧允许 ESX/ESXi 将较大的帧发送到物理网络上。要使巨帧生效，网络必须端到端支持巨帧。最大可支持 9 KB（9000 字节）的巨帧。

表 2-1 说明 ESX/ESXi 提供的对巨帧的支持级别。

表 2–1 对巨帧的支持

iSCSI 适配器的类型	巨帧支持	备注
软件 iSCSI	支持	与供应商核对，以确保您的物理网络适配器和 iSCSI 存储器支持巨帧。
从属硬件 iSCSI	支持	Broadcom 从属 iSCSI 适配器不支持巨帧。
独立硬件 iSCSI	不支持	

为软件和从属硬件 iSCSI 启用巨帧

使用 vSphere CLI 对指定用于 iSCSI 流量的每个 vSwitch 启用巨帧。同样，您必须创建启用巨帧的 VMkernel 网络接口。

1 在 vSwitch 上启用巨帧第 36 页，

要在 vSwitch 上启用巨帧，可更改该 vSwitch 的 MTU 配置。

2 创建已启用巨帧的 VMkernel 接口第 36 页，

使用 vSphere CLI 可创建启用了巨帧的 VMkernel 网络接口。

在 vSwitch 上启用巨帧

要在 vSwitch 上启用巨帧，可更改该 vSwitch 的 MTU 配置。

步骤

1 在 vSphere CLI 中，运行 **vicfg-vswitch -m MTU *vSwitch*** 命令。

通过此命令，可为此 vSwitch 上的所有物理网卡设置 MTU。所设置的 MTU 大小应在连接 vSwitch 的所有网卡中是最大的。

2 运行 **vicfg-vswitch -l** 命令可以在主机上显示 vSwitch 列表并检查 vSwitch 的配置是否正确。

创建已启用巨帧的 VMkernel 接口

使用 vSphere CLI 可创建启用了巨帧的 VMkernel 网络接口。

步骤

1 在 vSphere CLI 上，运行 `vicfg-vmknic` 命令创建支持巨帧的 VMkernel 连接。

vicfg-vmknic -a -I *ip address* -n *netmask* -m *MTU port group name*

2 运行 `vicfg-vmknic -l` 命令可显示 VMkernel 接口列表并检查启用了巨帧的接口的配置是否正确。

3 检查 VMkernel 接口是否已连接启用巨帧的 vSwitch。

4 将所有的物理交换机以及与该 VMkernel 接口相连的任何物理机或虚拟机配置为支持巨帧。

配置 iSCSI 启动器的发现地址

设置目标发现地址，以便 iSCSI 启动器确定网络上可供访问的存储资源。

ESX/ESXi 系统支持以下发现方法：

动态发现

也称为 SendTargets 发现。启动器每次与指定的 iSCSI 服务器联系时，都会向该服务器发送 SendTargets 请求。服务器通过向启动器提供一个可用目标的列表来做出响应。这些目标的名称和 IP 地址显示在**静态发现**选项卡上。如果移除了通过动态发现添加的静态目标，则该目标可在下次进行重新扫描、重置 HBA 或重新引导主机时返回到列表中。

静态发现

启动器不必执行任何发现。启动器拥有它可以联系的目标列表，并使用目标的 IP 地址和名称与这些目标进行通信。

设置动态发现

使用动态发现时，启动器每次与指定的 iSCSI 服务器联系时，都会向该服务器发送 SendTargets 请求。服务器通过向启动器提供一个可用目标的列表来做出响应。

设置动态发现时，只能添加新的 iSCSI 服务器。不能更改现有 iSCSI 服务器的 IP 地址、DNS 名称或端口号。要进行更改，请删除现有服务器，并添加一台新服务器。

步骤

1 登录 vSphere Client，在“清单”面板中选择服务器。
2 单击**配置**选项卡，然后在“硬件”面板中单击**存储适配器**。

   此时将显示可用存储适配器的列表。
3 选择要配置的 iSCSI 启动器，然后单击**属性**。
4 在“iSCSI 启动器属性”对话框中，单击**动态发现**选项卡。
5 要为 SendTargets 发现添加地址，请单击**添加**。

   此时将显示**添加发送目标服务器**对话框。
6 输入存储系统的 IP 地址或 DNS 名称，然后单击**确定**。

   在主机与此系统建立 SendTargets 会话后，新发现的任何目标均将出现在“静态发现”列表中。
7 要删除特定的发送目标服务器，请将其选中，然后单击**移除**。

   在移除发送目标服务器之后，它可能仍作为静态目标的父目标出现在“继承”字段中。此条目表示静态目标的发现位置且并不影响功能。

下一步

为 iSCSI 适配器配置动态发现后，请重新扫描适配器。

设置静态发现

借助于 iSCSI 启动器以及动态发现方法，可以使用静态发现并手动输入目标的信息。

设置静态发现时，只能添加 iSCSI 目标。不能更改现有目标的 IP 地址、DNS 名称、iSCSI 目标名称或端口号。要进行更改，请移除现有目标，然后添加一个新目标。

步骤

1 登录 vSphere Client，在“清单”面板中选择服务器。

2 单击**配置**选项卡，然后在“硬件”面板中单击**存储适配器**。

此时将显示可用存储适配器的列表。

3 选择要配置的 iSCSI 启动器，然后单击**属性**。

4 在“iSCSI 启动器属性”对话框中，单击**静态发现**选项卡。

该选项卡将显示所有动态发现目标和已输入的所有静态目标。

5 要添加目标，请单击**添加**，然后输入目标的信息。

6 要删除特定目标，请选择目标，然后单击**移除**。

下一步

为 iSCSI 适配器配置静态发现后，请重新扫描适配器。

配置 iSCSI 适配器的 CHAP 参数

由于 iSCSI 技术用来连接远程目标的 IP 网络不保护其传输的数据，因此必须确保连接的安全。质询握手身份验证协议 (CHAP) 是 iSCSI 实现的协议之一，该协议会验证访问网络上目标的启动器的合法性。

在主机和目标建立连接时，CHAP 使用三路握手算法验证主机和 iSCSI 目标（如果适用的话）的身份。系统根据启动器和目标共享的预定义的专用值或 CHAP 密钥进行验证。

ESX/ESXi 支持适配器级别的 CHAP 身份验证。在这种情况下，所有目标从 iSCSI 启动器接收相同的 CHAP 名称和密钥。对于软件和从属硬件 iSCSI 适配器，ESX/ESXi 还支持基于每个目标的 CHAP 身份验证，此身份验证使您能够为每个目标配置不同凭据以实现更高级别的安全性。

选择 CHAP 身份验证方法

ESX/ESXi 支持为所有类型的 iSCSI 启动器设置单向 CHAP，以及为软件和从属硬件 iSCSI 设置双向 CHAP。

在配置 CHAP 前，请检查是否在 iSCSI 存储系统中启用了 CHAP 并检查系统所支持的 CHAP 身份验证方法。如果已启用 CHAP，请为启动器启用 CHAP，并确保 CHAP 身份验证凭据与 iSCSI 存储器上的身份验证凭据相匹配。

ESX/ESXi 支持下列 CHAP 身份验证方法：

单向 CHAP 在单向 CHAP 身份验证中，目标需验证启动器，但启动器无需验证目标。

双向 CHAP 在双向 CHAP 身份验证中，提供了供启动器验证目标的附加安全级别。VMware 仅对软件和从属硬件 iSCSI 适配器支持此方法。

对于软件和从属硬件 iSCSI 适配器，可以为每个启动器或在目标级别设置单向 CHAP 和双向 CHAP。硬件 iSCSI 仅支持启动器级别的 CHAP。

设置 CHAP 参数时，请指定 CHAP 的安全级别。

注意 指定 CHAP 安全级别时，存储阵列的响应方式取决于阵列的 CHAP 实施，且因供应商而异。例如，选择使用 CHAP，除非已被目标禁止时，有些存储阵列在响应时会使用 CHAP，而有些则不会。有关不同的启动器和目标配置中的 CHAP 身份验证行为的信息，请参阅阵列文档。

表 2–2 CHAP 安全级别

CHAP 安全级别	描述	受支持的
不使用 CHAP	主机不使用 CHAP 身份验证。如果当前已启用，则选择此选项禁用身份验证。	软件 iSCSI 从属硬件 iSCSI 独立硬件 iSCSI
不使用 CHAP，除非目标需要	主机首选非 CHAP 连接，但如果目标要求可以使用CHAP 连接。	软件 iSCSI 从属硬件 iSCSI
使用 CHAP，除非已被目标禁止	主机首选 CHAP，但如果目标不支持 CHAP，可以使用非 CHAP 连接。	软件 iSCSI 从属硬件 iSCSI 独立硬件 iSCSI
使用 CHAP	主机需要成功的 CHAP 身份验证。如果 CHAP 协商失败，则连接失败。	软件 iSCSI 从属硬件 iSCSI

设置 iSCSI 启动器的 CHAP 凭据

您可以在启动器级别将所有目标设置为从 iSCSI 启动器接收相同的 CHAP 名称和密钥。默认情况下，所有发现地址或静态目标都继承在启动器级别设置的 CHAP 参数。

前提条件

在设置软件或从属硬件 iSCSI 的 CHAP 参数之前，请先确定是要配置单向 CHAP 还是双向 CHAP。独立硬件 iSCSI 适配器不支持双向 CHAP。

- 在单向 CHAP 中，目标会验证启动器。
- 在双向 CHAP 中，目标和启动器会相互验证。确保对 CHAP 和双向 CHAP 使用不同的密钥。

在配置 CHAP 参数时，确保它们与存储器端的参数相匹配。

CHAP 名称不得超过 511 个字母数字字符，CHAP 密钥不得超过 255 个字母数字字符。有些适配器（例如 QLogic 适配器）的限值更低，CHAP 名称不得超过 255 个字母数字字符，CHAP 密钥不得超过 100 个字母数字字符。

步骤

1 访问“iSCSI 启动器属性”对话框。

2 在**常规**选项卡上，单击 **CHAP**。

3 要配置单向 CHAP，请在 CHAP 下指定以下项：

   a 选择 CHAP 安全级别。

      - **不使用 CHAP，除非目标需要**（仅限软件 iSCSI 和从属硬件 iSCSI）
      - **使用 CHAP，除非已被目标禁止**
      - **使用 CHAP**（仅限软件 iSCSI 和从属硬件 iSCSI）。要能够配置双向 CHAP，必须选择此选项。

   b 指定 CHAP 名称。

      确保指定的名称与在存储器端配置的名称相匹配。

      - 要将 CHAP 名称设置为 iSCSI 启动器名称，请选中**使用启动器名称**。
      - 要将 CHAP 名称设置为除 iSCSI 启动器名称之外的任何其他名称，请取消选中**使用启动器名称**，并在**名称**字段中输入名称。

   c 输入单向 CHAP 密钥以用作身份验证的一部分。确保使用在存储器端输入的相同密钥。

4 要配置双向 CHAP，请先按照步骤 3 中的说明配置单向 CHAP。

   确保为单向 CHAP 选择**使用 CHAP** 选项。然后，在**双向 CHAP** 下，指定以下各项：

   a 选择**使用 CHAP**。

   b 指定双向 CHAP 名称。

   c 输入双向 CHAP 密钥。确保对单向 CHAP 和双向 CHAP 使用不同的密钥。

5 单击**确定**。

6 重新扫描启动器。

如果更改了 CHAP 或双向 CHAP 参数，则它们会用于新的 iSCSI 会话。但对于现有会话，直到注销并重新登录之后才能使用新设置。

设置目标的 CHAP 凭据

对于软件和从属硬件 iSCSI 适配器，可以为每个发现地址或静态目标配置不同的 CHAP 凭据。

在配置 CHAP 参数时，确保它们与存储器端的参数相匹配。CHAP 名称不得超过 511 个字母数字字符，CHAP 密钥不得超过 255 个字母数字字符。

前提条件

在设置软件和从属硬件 iSCSI 的 CHAP 参数之前，请先确定是要配置单向 CHAP 还是双向 CHAP。

- 在单向 CHAP 中，目标会验证启动器。
- 在双向 CHAP 中，目标和启动器会相互验证。确保对 CHAP 和双向 CHAP 使用不同的密钥。

步骤

1 访问“iSCSI 启动器属性”对话框。

2 选择**动态发现**或**静态发现**选项卡。

3 从可用目标的列表中，选择要配置的目标，然后单击**设置 > CHAP**。

4 在 CHAP 区域配置单向 CHAP。

   a 取消选中**从父项继承**。

   b 选择下列选项之一：

      - **不使用 CHAP，除非目标需要**
      - **使用 CHAP，除非已被目标禁止**
      - **使用 CHAP**。要能够配置双向 CHAP，必须选择此选项。

   c 指定 CHAP 名称。

      确保指定的名称与在存储器端配置的名称相匹配。

      - 要将 CHAP 名称设置为 iSCSI 启动器名称，请选中**使用启动器名称**。
      - 要将 CHAP 名称设置为除 iSCSI 启动器名称之外的任何其他名称，请取消选中**使用启动器名称**，并在**名称**字段中输入名称。

   d 输入单向 CHAP 密钥以用作身份验证的一部分。确保使用在存储器端输入的相同密钥。

5 要配置双向 CHAP，请先按照步骤 4 中的说明配置单向 CHAP。

确保为单向 CHAP 选择**使用 CHAP** 选项。然后，在“双向 CHAP”区域指定以下项：

a 取消选中**从父项继承**。

b 选择**使用 CHAP**。

c 指定双向 CHAP 名称。

d 输入双向 CHAP 密钥。确保对单向 CHAP 和双向 CHAP 使用不同的密钥。

6 单击**确定**。

7 重新扫描启动器。

如果更改了 CHAP 或双向 CHAP 参数，则它们会用于新的 iSCSI 会话。但对于现有会话，直到注销并重新登录之后才能使用新设置。

禁用 CHAP

如果存储系统不需要 CHAP，则可以将其禁用。

如果在需要 CHAP 身份验证的系统上禁用 CHAP，则现有 iSCSI 会话会保持活动状态，直到重新引导 ESX/ESXi 主机或者存储系统强制注销为止。在会话结束之后，您将不能再连接需要 CHAP 的目标。

步骤

1 打开“CHAP 凭据”对话框。

2 对于软件和从属硬件 iSCSI 适配器，若要仅禁用双向 CHAP 而保留单向 CHAP，请在“双向 CHAP”区域选择**不使用 CHAP**。

3 要禁用单向 CHAP，请在 CHAP 区域中选择**不使用 CHAP**。

如果设置了双向 CHAP，则在禁用单向 CHAP 时，双向 CHAP 将自动转换为**不使用 CHAP**。

4 单击**确定**。

配置 iSCSI 的其他参数

可能需要为 iSCSI 启动器配置其他参数。例如，有些 iSCSI 存储系统要求 ARP（地址解析协议）重定向，以在端口间动态移动 iSCSI 流量。在这种情况下，必须在主机上激活 ARP 重定向。

表 2-3 列出了使用 vSphere Client 可以配置的高级 iSCSI 参数。此外，可以使用 `vicfg-iscsi` vSphere CLI 命令配置部分高级参数。有关信息，请参阅《vSphere 命令行界面安装和脚本指南及 vSphere 命令行界面参考》。

除非在与 VMware 支持团队进行合作，或拥有为设置所提供值的全面信息，否则不要对高级 iSCSI 设置进行任何更改。

表 2-3 iSCSI 启动器的其他参数

高级参数	描述	配置对象
头摘要	增加数据完整性。启用头摘要后，系统会对每个 iSCSI 协议数据单元 (PDU) 的头部计算校验和，并使用 CRC32C 算法进行验证。	软件 iSCSI 从属硬件 iSCSI
数据摘要	增加数据完整性。启用数据摘要后，系统会对每个 PDU 的数据部分计算校验和，并使用 CRC32C 算法进行验证。 **注意** 使用 Intel Nehalem 处理器的系统会清除软件 iSCSI 的 iSCSI 摘要计算，因此可以减少对性能的影响。	软件 iSCSI 从属硬件 iSCSI
最大未完成 R2T 数	定义在收到确认 PDU 前可转换的 R2T（即将传输）PDU。	软件 iSCSI 从属硬件 iSCSI

表 2–3 iSCSI 启动器的其他参数 （续）

高级参数	描述	配置对象
初次突发长度	指定在执行单个 SCSI 命令期间 iSCSI 启动器可以发送到目标的未经请求的数据的最大数量，以字节为单位。	软件 iSCSI 从属硬件 iSCSI
最大突发长度	传入数据或请求的传出数据 iSCSI 序列中的最大 SCSI 数据负载，以字节为单位。	软件 iSCSI 从属硬件 iSCSI
最大接收数据段长度	在 iSCSI PDU 中可以接收的最大数据段长度，以字节为单位。	软件 iSCSI 独立硬件 iSCSI
会话恢复超时	指定执行会话恢复的时间限制（秒）。如果该超时时间超过限值，则 iSCSI 启动器将终止会话。	软件 iSCSI 从属硬件 iSCSI
无操作时间间隔	指定在从 iSCSI 启动器发送到 iSCSI 目标的 NOP-Out 请求之间的时间间隔（秒）。NOP-Out 请求充当验证 iSCSI 启动器和 iSCSI 目标之间的连接是否处于活动状态的 ping 机制。	软件 iSCSI 从属硬件 iSCSI
无操作超时	指定主机收到 NOP-In 消息之前的时间限制（秒）。该消息由 iSCSI 目标发出，以响应 NOP-Out 请求。无操作超时时间超过限值时，启动器将终止当前会话并启动新的会话。	软件 iSCSI 从属硬件 iSCSI
ARP 重定向	允许存储系统将 iSCSI 流量从一个端口动态移动到另一个端口。ARP 对执行基于阵列的故障切换的存储系统是必需的。	软件和独立硬件 iSCSI （可通过 vSphere CLI 配置）
延迟的 ACK	允许系统对接收的数据包进行延迟确认。	软件 iSCSI 从属硬件 iSCSI

配置 iSCSI 的高级参数

高级 iSCSI 设置控制如标头、数据摘要、ARP 重定向、延迟的 ACK 等参数。通常，不需要更改这些设置，因为 ESX/ESXi 主机使用分配的预定义值能够正常运行。



小心 除非在与 VMware 支持团队进行合作，或拥有为设置所提供值的全面信息，否则不要对高级 iSCSI 设置进行任何更改。

步骤

1 访问“iSCSI 启动器属性”对话框。

2 要在启动器级别配置高级参数，请在“常规”选项卡上单击**高级**。继续步骤 4。

3 在目标级别配置高级参数。

   在目标级别，只能为软件和从属硬件 iSCSI 适配器配置高级参数。

   a 选择**动态发现**或**静态发现**选项卡。

   b 从可用目标的列表中，选择要配置的目标，然后单击**设置 > 高级**。

4 为要修改的高级参数输入任何所需值，然后单击**确定**保存更改。

iSCSI 会话管理

要相互通信，iSCSI 启动器和目标需要建立 iSCSI 会话。您可以使用 vSphere CLI 查看和管理 iSCSI 会话。

默认情况下，软件 iSCSI 和从属硬件 iSCSI 启动器会在每个启动器端口与每个目标端口之间启动一个 iSCSI 会话。如果 iSCSI 启动器或目标有多个端口，则主机可以建立多个会话。每个目标的默认会话数等于 iSCSI 适配器上的端口数乘以目标端口数。您可以显示所有当前会话，以便对它们进行分析和调试。

要创建多条指向存储系统的路径，可以通过复制 iSCSI 适配器与目标端口之间的现有会话来增大默认会话数。

此外，也可以建立与特定目标端口的会话。如果主机连接到默认情况下为启动器提供一个目标端口，但可将其他会话重定向到不同目标端口的单端口存储系统，则此操作很有用。在 iSCSI 启动器与另一个目标端口之间建立新会话，将创建一条指向存储系统的其他路径。



小心　某些存储系统不支持源自同一启动器名称或端点的多个会话。尝试创建多个与此类目标的会话可能会导致 iSCSI 环境的行为不可预知。

查看 iSCSI 会话

使用 vSphere CLI 可显示 iSCSI 适配器与存储系统之间的 iSCSI 会话。

步骤

1　要列出适配器的会话，使用 **esxcli swiscsi session list -d *vmhbaXX***。

vmhbaXX 是 iSCSI 适配器的名称。

2　要列出目标的会话，请使用 **esxcli swiscsi session list -t *iqnXX* -d *vmhbaXX***。

iqnXX 是 iSCSI 适配器连接至的目标的名称。

示例 2-3　查看指定目标的会话

```
# esxcli swiscsi session list -t iqn.2001-05.com.equallogic:0-8a0906-cbcde0d02-6a445baadf44ab92-
vcli7 -d vmhba36
```

为指定目标添加 iSCSI 会话

使用 vSphere CLI 可为指定的目标添加 iSCSI 会话。

步骤

1　要在 iSCSI 适配器级别添加会话，使用 **esxcli swiscsi session add -d *vmhbaXX***。

vmhbaXX 是 iSCSI 适配器的名称。

2　要在目标级别添加会话，使用 **esxcli swiscsi session add -t *iqnXX* -d *vmhbaXX***。

iqnXX 是 iSCSI 适配器连接至的目标的名称。

示例 2-4　为指定目标添加 iSCSI 会话

```
# esxcli swiscsi session add -t iqn.2001-05.com.equallogic:0-8a0906-42f9b0401-7d40d28af2a4ade1-
vcli7 -d vmhba36
```

下一步

重新扫描 iSCSI 适配器。

复制 iSCSI 会话

使用 vSphere CLI 可复制 iSCSI 适配器与目标之间的 iSCSI 会话。通过复制会话，可增大默认会话数并创建指向存储系统的其他路径。

步骤

◆　要复制会话，使用 **esxcli swiscsi session add -d *vmhbaXX* -t *iqnXX* -s *session_isid***。

- d - iSCSI 适配器的名称，如 *vmhbaXX*。
- t - 要登录的目标的名称，如 *iqnXX*。
- s - 要复制的会话的 ISID，如 *session_isid*。您可以通过列出所有会话来进行查找。

示例 2-5 复制 iSCSI 会话

```
# esxcli swiscsi session add -t iqn.2001-05.com.equallogic:0-8a0906-cbcde0d02-6a445baadf44ab92-
vcli7 -s '00:02:3d:00:00:01' -d vmhba36
```

下一步

重新扫描 iSCSI 适配器。

删除 iSCSI 会话

使用 vSphere CLI 可删除 iSCSI 适配器与目标之间的 iSCSI 会话。

步骤

◆ 要移除会话，使用 **esxcli swiscsi session remove**。该命令采用以下选项。

  - d - iSCSI 适配器的名称，如 *vmhbaXX*。
  - t - 要登录的目标的名称，如 *iqnXX*。
  - s - 要复制的会话的 ISID，如 *session_isid*。您可以通过列出所有会话来进行查找。

示例 2-6 删除 iSCSI 会话

```
# esxcli swiscsi session remove -t iqn.2001-05.com.equallogic:0-8a0906-cbcde0d02-6a445baadf44ab92-
vcli7 -s '00:02:3d:01:00:01'
```

下一步

重新扫描 iSCSI 适配器。

添加 iSCSI 存储器

在 iSCSI 存储设备上创建数据存储时，添加存储器向导将指导您完成所有配置。

步骤

1 登录 vSphere Client，在“清单”面板中选择服务器。
2 单击**配置**选项卡，然后单击**存储器**。
3 单击**添加存储器**。
4 选择**磁盘/LUN** 存储器类型，然后单击**下一步**。

   此时将显示**选择磁盘/LUN** 页面。这可能要花几秒钟时间，具体视目标数而定。
5 选择要用于数据存储的 iSCSI 设备，然后单击**下一步**。
6 检查当前磁盘布局，然后单击**下一步**。
7 输入数据存储名称，然后单击**下一步**。

   数据存储名称将显示在 vSphere Client 中，且标签在当前 VMware vSphere 实例中必须是唯一的。
8 如果需要的话，请调整用于数据存储的文件系统值和容量。

   默认情况下，存储设备上的全部可用空间均可供使用。
9 单击**下一步**。

   此时将显示**即将完成**页面。
10 检查数据存储配置信息，然后单击**完成**。

数据存储现在可以在 iSCSI 存储设备上使用。

为 ESX/ESXi 修改 SAN 存储系统 3

配置 iSCSI 启动器和存储器之后，可能需要修改存储系统以确保其能够与 ESX/ESXi 实施正确地配合使用。

本节将论述支持与 VMware ESX/ESXi 配合使用的几种 iSCSI 存储系统。对于每种设备，本章都将列出主要的已知潜在问题，提供特定于供应商的信息（如果有），或附加 VMware 知识库文章中的信息。

注意 本节中的信息仅随各版本更新。新的信息或许已可通过其他途径获取。此外，还有一些其他的 iSCSI 存储系统也受支持，但本章未对它们进行论述。请参阅最新的《存储器/SAN 兼容性指南》，咨询存储供应商，并浏览 VMware 知识库文章。

本章讨论了以下主题：

测试 ESX/ESXi SAN 配置

ESX/ESXi 支持各种具有不同配置的 SAN 存储系统。通常，VMware 使用支持的存储系统测试 ESX/ESXi，测试内容包括基本连接、HBA 故障切换等。

并非所有存储设备都针对 ESX/ESXi 的所有特性和功能进行了认证，供应商可能就 ESX/ESXi 提供特定形式的支持。

VMware 会对使用存储系统的 ESX/ESXi 进行以下配置测试：

基本连接 测试 ESX/ESXi 能否识别存储系统并与存储系统共同运行。此配置不允许应用多路径或任何类型的故障切换。

iSCSI 故障切换 服务器配备多个 iSCSI HBA 或网卡。服务器在出现 HBA 或网卡故障时会保证正常运行。

存储器端口故障切换	服务器与多个存储器端口相连，并在发生存储器端口故障和交换机故障时能正常运行。
从 SAN 引导（仅限 ESX 主机）	ESX 主机从 SAN 上配置的 LUN 引导而不是从服务器自身引导。

iSCSI SAN 存储系统的常规注意事项

当准备与 ESX/ESXi 配合使用的 iSCSI SAN 存储系统时，需要遵循适用于所有存储系统的特定常规要求。

对于所有存储系统，存在以下常规要求：

- LUN 对每台主机的每个 HBA 呈现的 LUN ID 号都必须相同。如果使用了不同的号码，则 ESX/ESXi 主机便识别不出指向同一 LUN 的不同路径。有关如何配置相同 SAN LUN ID 的具体说明取决于供应商，请参阅存储器文档以获取详细信息。
- 除非为本章中所述的单个存储系统指定了主机类型，否则应将呈现给 ESX/ESXi 的 LUN 的主机类型设置为 `Linux` 或 `Linux Cluster`（如果适用于您的存储系统）。ESX/ESXi 用来访问存储系统的方法与 Linux 访问最兼容，但根据您所使用的存储系统，这种情况可能会有所不同。
- 如果使用了 vMotion、DRS 或 HA，则请确保虚拟机的源主机与目标主机都能看到具有相同 LUN ID 的同一 LUN。出于对可能出现数据损坏的担心，SAN 管理员可能认为使多个主机看见相同的 LUN 不合常规。但是，VMFS 可防止多个虚拟机同时对同一文件进行写入，因此给所有需要的 ESX/ESXi 系统置备 LUN 是可行的。
- 如果没有为正在访问的 LUN 设置 CHAP 身份验证，则必须同时禁用 ESX/ESXi 主机上的 CHAP。否则，存储系统的身份验证会失败，尽管 LUN 没有 CHAP 要求。

EMC CLARiiON 存储系统

EMC CLARiiON 存储系统可与采用 iSCSI SAN 配置的 ESX/ESXi 主机配合使用。通常，使用 EMC 软件执行配置操作。

有关详细信息，请参见 EMC 文档。

这是主动-被动磁盘阵列，因此适用于所有主动-被动磁盘阵列的任何有关问题都是相关的。此外，请记住以下几点：

- 为避免发生路径抖动的可能，默认多路径策略为“最近使用”，而非“固定”。ESX/ESXi 系统在识别存储系统时设置默认策略。
- 要从 SAN 引导，请在 HBA BIOS 中选择主动存储处理器作为引导 LUN 的目标。
- 在 EMC CLARiiON AX100i 和 AX150i 系统上，只有在使用 Navisphere Management Suite 管理 SAN 时才支持 RDM。Navisphere Express 不能保证可以正确配置 RDM。

  要成功使用 RDM，给定的 LUN 必须对群集中所有 ESX/ESXi 主机呈现相同的 LUN ID。默认情况下，AX100i 和 AX150i 不执行此操作。
- 使用 AX100i 或 AX150i 存储系统时，并无主机代理定期检查主机配置和将更改应用于存储系统。`axnaviserverutil cli` 实用程序将用于更新所作的更改。这是一项手动操作，应根据需要执行。
- EMC CLARiiON 存储系统上的端口绑定支持需要不同子网中的启动器。有关其他详细信息，请参见供应商文档。
- 要使 ESX/ESXi 支持使用 ALUA 的 EMC CLARiiON，请检查 HCL 以确保在存储阵列上使用正确的固件版本。有关其他信息，请与存储器供应商联系。

EMC Symmetrix 存储系统

要使用 ESX/ESXi，EMC Symmetrix 存储系统需要某些特定设置。使用 EMC 软件配置存储系统。有关信息，请参见 EMC 文档。

要在 Symmetrix 网络存储系统上执行 ESX/ESXi 操作，需要以下设置：

- 公共序列号 (C)
- 启用自动协商 (EAN)
- SCSI 3 (SC3) 集（已启用）
- 唯一的全球名称 (UWN)
- 需要 SPC-2 (Decal) (SPC2) SPC-2 标记

注意 ESX/ESXi 主机将来自 Symmetrix 存储系统、容量不超过 50 MB 的任何 LUN 均视为管理 LUN。这些 LUN 也称为伪 LUN 或网关 LUN。这些 LUN 可显示在 EMC Symmetrix Management Interface 中，并且不应用于保存数据。

使 HP StorageWorks MSA1510i 能够与 ESX/ESXi 通信

要使 HP StorageWorks MSA1510i 存储系统能够与 ESX/ESXi 主机进行通信，必须完成一些设置和配置步骤，本节将介绍这些步骤。

步骤

1 按照供应商安装文档中的详细说明安装、连接和启动网络设备。
2 获取分配给 MSA1510i 控制器管理端口的 IP 地址。
   a 滚动查看液晶面板上的消息，直到显示下列消息：`603 Port MA0 IP` *`address`*
   b 记录 **MSA1510i 基本信息**中显示的管理端口 IP 地址。
3 在 MSA1510i LAN 网段上的服务器或工作站中，打开 Web 浏览器，然后输入从步骤 2 中获取的地址。
4 出现提示时，输入默认访问权限。
   - 用户名：`root`
   - 密码：`root`
5 出现提示时，设置唯一的用户名和密码。
6 使用向导完成以下操作。

选项	描述
存储器配置	a 设置容错模式（RAID 模式）。 b 根据适当的 RAID 级别分配备用磁盘。
iSCSI 配置（配置 iSCSI 入口）	a 选择数据端口。 b 为数据端口分配 IP 地址。 c VLAN 在交换机上设置，并用作控制访问存储器的一种方法。如果使用 VLAN，请输入要使用的 VLAN ID（0 = 不使用）。 d 向导建议使用默认 iSCSI 目标名称和 iSCSI 目标别名。接受默认值或输入用户定义的值。 **注意** 要配置其余数据端口，请完成初始系统配置向导过程，然后使用**配置**选项卡上的可用任务。

选项	描述
登录设置	输入登录设置。
管理设置	输入管理设置。

7 单击**完成**以应用配置设置。

注意 向导仅对于基本配置任务可用。在**管理**和**配置**选项卡中可查看和更改配置。

下一步

完成初始设置后，执行以下任务来完成配置：

- 创建阵列。
- 创建逻辑驱动器。
- 创建目标。
- 创建入口组。
- 将使用向导创建的入口与创建的入口组关联起来，或将入口分配到入口组。
- 将逻辑驱动器映射到目标。
- 添加启动器（启动器 IQN 名称和别名）。
- 更新逻辑驱动器的 ACL 以提供对启动器的访问（选择启动器列表以访问逻辑驱动器）。

HP StorageWorks EVA 存储系统

HP StorageWorks EVA 系统的两种类型分别为 EVA_GL 和 EVA_XL，前者为主动-被动系统，后者为主动-主动系统。对于要与 ESX/ESXi 配合使用的系统，某些特定设置是必需的。

向 ESX/ESXi 主机显示 LUN 时请将连接类型设置为**自定义**。请采用以下值之一：

- 对于 HP EVAgl 3000/5000（主动-被动），请使用主机模式类型 `000000002200282E` 。
- 对于 HP EVAgl 固件 4.001（GL 系列的主动-主动固件）及更高版本，使用主机模式类型 `VMware`。
- 对于固件版本低于 5.031 的 EVA4000/6000/8000 主动-主动阵列，请使用主机模式类型 `000000202200083E`。
- 对于固件版本不低于 5.031 的 EVA4000/6000/8000 主动-主动阵列，请使用主机模式类型 `VMware`。

对于其他情况，EVA 系统无需特殊的配置更改即可与 ESX/ESXi 系统配合使用。

NetApp 存储系统

要使 NetApp 存储系统能够在 ESX/ESXi 环境中进行通信，必须执行特定的配置步骤。

有关 NetApp 和 VMware 最佳实践和 SAN 解决方案的其他文档，请搜索 NetApp 网页。

表 3–1 配置步骤

配置步骤	描述
禁用 ALUA。	如果任何 iSCSI 启动器是启动器组 (igroup) 的一部分，则在 NetApp 筛选器上禁用 ALUA。
设置多路径。	在两个 iSCSI HBA 之间设置多路径及在 NetApp 存储系统上设置多个端口时，为这两个 HBA 提供不同的动态或静态发现地址以连接存储器。 NetApp 存储系统仅允许为每个目标和每个启动器建立一个连接。试图建立更多连接会导致第一个连接断开。因此，单个 HBA 不应尝试连接与同一 NetApp 目标关联的多个 IP 地址。

表 3–1 配置步骤 （续）

配置步骤	描述
设置 LUN 类型和启动器组类型。	为存储系统设置适当的 LUN 类型和启动器组类型： ■ LUN 类型 - VMware（如果 VMware 类型不可用，请使用 Linux）。 ■ 启动器组类型 - VMware（如果 VMware 类型不可用，请使用 Linux）。
置备存储器。	使用 FilerView 或 CLI。

使用 FilerView 存储器管理置备存储器

可以使用 FilerView 在 NetApp 存储系统上置备存储器。

步骤

1 登录 NetApp 存储系统管理 (FilerView)。

2 创建卷。

a 选择**卷**并单击**添加**。

b 单击**下一步**。

c 选择**灵活性**（默认）或**传统**，然后单击**下一步**。

d 输入**卷名称**，选择一种**语言**，然后单击**下一步**。

e 输入**容量总计**、**卷大小总计**和**空间保证**的值，然后单击**下一步**。

f 单击**提交**创建卷。

3 创建 LUN。

a 选择 **LUN**，然后单击**添加**。

b 输入以下内容：

- **路径**：输入路径，例如 `/vol/vol1/lun1`。
- **LUN 协议类型**：VMware。
- **描述**：简要描述。
- **大小和单位**：输入大小（如 10 GB），然后选择**保留空间**。

4 创建启动器组。

   a 选择 **LUN > 启动器组**，并单击**添加**。

   b 输入以下内容：

      - **组名称**：输入组名称
      - **类型**：选择 **iSCSI**。
      - **操作系统**：输入 `VMware`。
      - **启动器**：输入完全限定启动器名称。如果有多个启动器，则必须使用回车分隔每个启动器。

   c 单击**添加**。

5 将 LUN 映射至启动器组。

   a 选择 **LUN** 并单击**管理**。

      此时将显示 LUN 列表。

   b 从该列表中，单击特定 LUN 的**映射**行上的标签。

   c 单击**添加要映射的组**。

   d 选择启动器组，并单击**添加**。

   e 出现提示时，输入 LUN ID（从 0 到 255 的任意数）并单击**应用**。

使用命令行界面置备存储器

可以使用命令行界面在 NetApp 存储系统上置备存储器。

步骤

1 根据需要使用命令行界面创建聚合。

   ```
   aggr create vmware-aggrnumber of disks
   ```

2 创建灵活卷。

   ```
   vol create aggregate namevolume size
   ```

3 创建 Qtree 存储各个 LUN。

   ```
   qtree create path
   ```

4 创建 LUN。

   ```
   lun create -s size -t vmware path
   ```

5 创建启动器组。

   ```
   igroup create -f -t vmware igroup name
   ```

6 将 LUN 映射至已创建的启动器组。

   ```
   lun map (path) igroup nameLUN ID
   ```

EqualLogic 存储系统

要对即将在 ESX/ESXi 实施中运行的 EqualLogic 存储系统进行设置，必须解决某些特定问题。

以下是 EqualLogic 存储系统与 ESX/ESXi 配合使用的特定要求：

- 多路径。无需进行特殊设置，因为 EqualLogic 存储系统支持对 iSCSI 透明的存储处理器故障切换。多个 iSCSI HBA 或网卡可连接存储器端的同一目标或 LUN。
- 创建 iSCSI LUN。从 EqualLogic Web 入口，右键单击**卷**，然后选择**创建卷**。
- 对 ESX/ESXi 硬件 iSCSI HBA 启用 ARP 重定向。
- EqualLogic 存储系统对每个存储池的 iSCSI 连接最大限制是 512，对每个存储组的 iSCSI 连接最大限制是 2048。

有关配置和使用 EqualLogic 存储系统的详细信息，请参见供应商文档。

LeftHand Networks SAN/iQ 存储系统

SAN/iQ SAN 支持从软件启动器和硬件启动器连接 ESX/ESXi iSCSI。

配置 SAN/iQ 时，为 SAN/iQ 存储设备启用自动卷再签名以允许访问 SAN/iQ 快照和远程副本。

有关为 VMware vSphere 配置 LeftHand Networks SAN 的详细信息，请参见与 VMware 有关的供应商文档。

基本配置步骤包括多个任务。

1 安装 SAN/iQ 存储节点。
2 创建 SAN/iQ 管理组和群集。
3 创建卷。
4 将卷分配给身份验证组和卷列表。
5 对硬件 iSCSI HBA 启用 ARP 重定向。

最好在 SAN/iQ 中为所有 ESX/ESXi 身份验证组配置虚拟 IP 负载平衡。

Dell PowerVault MD3000i 存储系统

这些特殊注意事项在为 MD3000i iSCSI 存储系统配置双向 CHAP 时适用。

为 MD3000i iSCSI 阵列配置双向 CHAP 时，请遵循以下准则：

- 在 MD3000i 存储系统上，双向 CHAP 配置仅需要 CHAP 密钥。
- 在 ESX/ESXi 主机上，双向 CHAP 配置需要名称和 CHAP 密钥。在 ESX/ESXi 主机上配置双向 CHAP 时，将目标的 IQN 名称作为双向 CHAP 名称输入。请确保 CHAP 密钥与在阵列上设置的密钥相匹配。

vApps 中的 iSCSI 目标

如果在虚拟设备（如 HP LeftHand P4000 VSA）中使用 iSCSI 目标，则主机应通过软件 iSCSI 适配器而不是硬件 iSCSI 适配器连接到该目标。

从 iSCSI SAN 引导 4

将主机设置为从 SAN 引导时，主机的引导映像存储在 SAN 存储系统中的一个或多个 LUN 上。主机在启动时，将从 SAN 上的 LUN 引导，而不是从其本地磁盘引导。

如果您不想处理本地存储器的维护，或者具有无磁盘硬件配置（如刀片系统），则可以使用从 SAN 引导。

ESX 和 ESXi 主机支持从 SAN 引导的不同方法。

表 4–1 从 iSCSI SAN 引导支持

主机类型	独立硬件 iSCSI	软件 iSCSI 和从属硬件 iSCSI
ESX 主机	受支持。从 SAN 引导需要 iSCSI HBA。	不支持。
ESXi 主机	不支持。	受支持。网络适配器必须支持 iBFT。

从 SAN 引导主机有许多优点。

- 服务器成本更低。无需内部存储器，可更密集地安置服务器，且运行过程散热也更少。
- 服务器更换更方便。可以更换服务器并将新服务器指向旧的引导位置。
- 减少了空间浪费。
- 备份过程更轻松。可作为 SAN 整体备份过程的一部分来备份 SAN 中的系统引导映像。
- 改善了管理。创建和管理操作系统映像变得更简单且更高效。

本章讨论了以下主题：

关于从 iSCSI SAN 引导的常规建议

如果您计划将 iSCSI LUN 设置为主机的引导设备并使用它，则您需要遵循某些常规准则。

这些准则适用于从软件和从属 iSCSI 引导以及从独立硬件 iSCSI 引导。

- 查看供应商针对引导配置中使用的硬件提出的任何建议。
- 有关安装必备条件和要求，请查看相应的 ESX/ESXi 安装指南。
- 使用静态 IP 地址以减少出现 DHCP 冲突的机会。

- 对 VMFS 数据存储和引导分区使用不同的 LUN。
- 在存储系统上配置正确的 ACL。
  - 引导 LUN 只应对使用该 LUN 的主机可见。不应允许 SAN 上的其他任何主机看到该引导 LUN。
  - 如果 LUN 用于 VMFS 数据存储，则可由多个主机共享。存储系统上的 ACL 可允许您这样做。
- 配置诊断分区。
  - 可以将诊断分区放置在引导 LUN 上（仅适用于独立硬件 iSCSI）。如果配置引导 LUN 中的诊断分区，则此 LUN 不能在多个主机中共享。如果某个单独的 LUN 用于诊断分区，则该 LUN 可由多个主机共享。
  - 如果使用 iBFT 从 SAN 引导 ESXi 主机，则不能在 SAN LUN 上设置诊断分区。而是应改用 vSphere Management Assistant (vMA) 从主机中收集诊断信息，并存储这些信息以进行分析。

    请参见第 60 页，“收集 ESXi 主机的诊断信息”。

准备 iSCSI SAN

在将 ESX/ESXi 主机配置为从 iSCSI LUN 引导之前，需要准备并配置存储区域网络。



小心 如果从 SAN 引导时使用脚本式安装来安装 ESX，则必须执行特殊步骤以避免数据意外丢失。

步骤

1 请参见适用于您的设置的任何线缆连接指南，连接网络线缆。

2 确保存储系统与服务器之间的 IP 连接的正确性。

  这包括正确配置存储网络上的任何路由器或交换机。存储系统必须能够 ping 主机中的 iSCSI 适配器。

3 配置存储系统。

  a 在存储系统上创建要从中引导主机的卷（或 LUN）。

  b 配置存储系统，以便主机能够访问所分配的 LUN。

    这可能涉及到用主机上所使用的 IP 地址、iSCSI 名称和 CHAP 身份验证参数来更新 ACL。在某些存储系统上，除了提供 ESX/ESXi 主机的访问信息外，还必须将分配的 LUN 与主机显式关联起来。

  c 确保已将 LUN 正确提供给主机。

  d 确保没有其他系统能够访问配置的 LUN。

  e 记下分配给主机的目标的 iSCSI 名称和 IP 地址。

    您必须使用此信息来配置 iSCSI 适配器。

将 ESX 主机配置为从 iSCSI SAN 引导

仅当 ESX 主机使用了独立硬件 iSCSI 适配器（例如 QLogic HBA）时，才能使用该主机从 iSCSI SAN 引导。需要配置独立硬件 iSCSI 适配器以从 SAN 引导。

此过程讨论如何启用 QLogic iSCSI HBA 以从 SAN 引导。有关详细信息以及有关 QLogic 适配器配置设置的最新详细信息，请参见 QLogic 网站。

前提条件

由于您先需要从 VMware 安装介质引导，因此，请将主机设置为从 DVD-ROM 引导。为此，请在系统 BIOS 设置中更改系统引导顺序。

步骤

1 将安装 DVD 插入到 DVD-ROM 驱动器中，然后重新引导主机。

2 先使用 BIOS 将主机设置为从 DVD-ROM 驱动器引导。

3 在服务器开机自检 (POST) 期间，按 Crtl+q 进入 QLogic iSCSI HBA 配置菜单。

4 选择要配置的 I/O 端口。

  默认情况下，适配器引导模式设置为“禁用”。

5 配置 HBA。

  a 在 **Fast!UTIL 选项**菜单中，选择**配置设置 > 主机适配器设置**。

  b 为主机适配器配置以下设置：启动器 IP 地址、子网掩码、网关、启动器 iSCSI 名称和 CHAP（如果需要）。

6 配置 iSCSI 设置。

  请参见第 55 页，“配置 iSCSI 引导设置”。

7 保存更改并重新启动系统。

配置 iSCSI 引导设置

将 ESX 主机设置为从 iSCSI 引导时，需要配置 iSCSI 引导设置。

步骤

1 在 **Fast!UTIL 选项** 菜单中，选择**配置设置 > iSCSI 引导设置**。

2 必须先将适配器引导模式设置为**手动**才能设置 SendTargets。

3 选择**主要引导设备设置**。

  a 输入发现**目标 IP** 和**目标端口**。

  b 如果指定地址只有一个要从其中引导的 iSCSI 目标和 LUN，则可将**引导 LUN** 和 **iSCSI 名称**字段留空。否则，必须指定这些字段以确保不会从某个其他系统的卷进行引导。到达目标存储系统之后，这些字段将在重新扫描后进行填充。

  c 保存更改。

4 在 **iSCSI 引导设置**菜单中，选择主要引导设备。此时将自动重新扫描 HBA 以查找新的目标 LUN。

5 选择 iSCSI 目标。

  注意 如果目标内存在多个 LUN，则可在找到 iSCSI 设备之后按 **Enter** 来选择特定 LUN ID。

6 返回至**主要引导设备设置**菜单。重新扫描后，**引导 LUN** 和 **iSCSI 名称**字段将被填充。将**引导 LUN** 的值更改为所需的 LUN ID。

iBFT iSCSI 引导概览

可以使用软件或从属硬件 iSCSI 适配器和网络适配器从 iSCSI SAN 引导 ESXi 主机。

要部署 ESXi 并从 iSCSI SAN 引导，主机必须具备可支持 iSCSI 引导固件表 (iBFT) 格式的能够进行 iSCSI 引导的网络适配器。iBFT 是一种将 iSCSI 引导设备的相关参数传递到操作系统的方法。

在安装 ESXi 并从 iSCSI SAN 引导之前，请先配置网络适配器上的网络参数和 iSCSI 引导参数，并启用此适配器以进行 iSCSI 引导。由于配置网络适配器因供应商而异，因此，请查看供应商文档来了解相关说明。

首次从 iSCSI 引导时，系统上的 iSCSI 引导固件将连接到 iSCSI 目标。如果登录成功，则固件会将网络和 iSCSI 引导参数保存到 iBFT，并将该表存储在系统内存中。系统可使用此表配置其自己的 iSCSI 连接和网络并启动。

以下列表介绍了 iBFT iSCSI 引导顺序。

1 重新启动后，系统 BIOS 将检测网络适配器上的 iSCSI 引导固件。

2 iSCSI 引导固件使用预先配置的引导参数与指定的 iSCSI 目标连接。

3 如果成功连接到 iSCSI 目标，则 iSCSI 引导固件会将网络参数和 iSCSI 引导参数写入 iBFT，并将该表存储在系统内存中。

注意 系统可使用此表配置其自己的 iSCSI 连接和网络并启动。

4 BIOS 将引导引导设备。

5 VMkernel 则开始加载并取代引导操作。

6 VMkernel 使用 iBFT 中的引导参数连接到 iSCSI 目标。

7 iSCSI 连接建立后，系统将进行引导。

iBFT iSCSI 引导注意事项

使用已启用 iBFT 的网络适配器从 iSCSI 引导 ESXi 主机时，有几点需要特别注意。

iBFT iSCSI 引导不支持以下各项：

- IPv6
- 为已启用 iBFT 的网络适配器进行故障切换

注意 在尝试安装和引导 VMware ESXi 4.1 版本之前，先使用供应商提供的工具更新网卡的引导代码和 iBFT 固件。请查看供应商文档和 VMware HCL 指南，获取 VMware ESXi 4.1 iBFT 引导支持的引导代码和 iBFT 固件版本。供应商在 ESXi 4.1 版本之前发布的引导代码和 iBFT 固件可能不起作用。

将主机设置为从 iBFT iSCSI 引导后，您必须注意以下限制：

- 不能禁用软件 iSCSI 适配器。如果 BIOS 中存在 iBFT 配置，则主机会在每次重新引导的过程中重新启用软件 iSCSI 适配器。
- 不能使用 vSphere Client 来移除 iBFT iSCSI 引导目标。该目标将在适配器静态目标列表中显示。

将 ESXi 主机设置为从 iSCSI 引导

对于 ESXi 主机，可以使用软件 iSCSI 适配器或者从属硬件 iSCSI 适配器和网络适配器从 iSCSI 引导。

将主机设置为从 iSCSI 引导时，需要执行多项任务。

1 配置 iSCSI 引导参数第 57 页，

要开始 iSCSI 引导过程，主机上的网络适配器必须具有专门配置的 iSCSI 引导固件。配置该固件时，需要指定网络参数和 iSCSI 参数，并启用适配器以进行 iSCSI 引导。

2 在 BIOS 中更改引导顺序第 57 页，

将主机设置为从 iBFT iSCSI 引导后，可以更改引导顺序来强制主机按相应的顺序进行引导。

3 将 ESXi 安装到 iSCSI 目标第 57 页，

将主机设置为从 iSCSI 引导时，需要将 ESXi 映像安装到目标 LUN。

4 从 iSCSI 目标引导 ESXi 第 58 页，

在准备好主机以进行 iSCSI 引导并将 ESXi 映像复制到 iSCSI 目标之后，可以执行实际引导。

配置 iSCSI 引导参数

要开始 iSCSI 引导过程，主机上的网络适配器必须具有专门配置的 iSCSI 引导固件。配置该固件时，需要指定网络参数和 iSCSI 参数，并启用适配器以进行 iSCSI 引导。

网络适配器上的配置可以是动态的，也可以是静态的。如果使用动态配置，则指示所有目标和启动器引导参数都将使用 DHCP 获得。对于静态配置，需要手动输入包括主机的 IP 地址和启动器 IQN 在内的数据以及目标参数。

步骤

◆ 在用于从 iSCSI 引导的网络适配器上，指定网络和 iSCSI 参数。

由于配置网络适配器因供应商而异，因此，请查看供应商文档来了解相关说明。

在 BIOS 中更改引导顺序

将主机设置为从 iBFT iSCSI 引导后，可以更改引导顺序来强制主机按相应的顺序进行引导。

对于 ESXi 安装，将 BIOS 引导顺序更改为以下顺序：

- iSCSI
- DVD-ROM

由于在 BIOS 中更改引导顺序因供应商而异，因此，请参考供应商文档来了解相关说明。以下示例步骤介绍了如何在使用 Broadcom 网络适配器的 Dell 主机上更改引导顺序。

步骤

1 打开该主机。
2 在开机自检 (POST) 期间，按 F2 进入 BIOS Setup。
3 在 BIOS Setup 中，选择 **Boot Sequence**，然后按 Enter。
4 在 Boot Sequence 菜单中，排列可引导项，以便 iSCSI 优先于 DVD-ROM。
5 按 Esc 退出 Boot Sequence 菜单。
6 按 Esc 退出 BIOS Setup。
7 选择 **Save Changes**，然后单击 **Exit** 退出 BIOS Setup 菜单。

将 ESXi 安装到 iSCSI 目标

将主机设置为从 iSCSI 引导时，需要将 ESXi 映像安装到目标 LUN。

前提条件

- 将引导网卡上的 iSCSI 引导固件配置为指向要用作引导 LUN 的目标 LUN。
- 在 BIOS 中更改引导顺序，以便 iSCSI 优先于 DVD-ROM。
- 如果使用 Broadcom 适配器，请将**引导至 iSCSI 目标**设置为**已禁用**。

步骤

1 将安装介质插入到 DVD-ROM 驱动器中，然后重新启动主机。
2 当安装程序启动时，遵循一般安装过程。
3 出现提示时，选择 iSCSI LUN 作为安装目标。

   安装程序将 ESXi 引导映像复制到 iSCSI LUN 中。
4 系统重新启动后，取出安装 DVD。

从 iSCSI 目标引导 ESXi

在准备好主机以进行 iSCSI 引导并将 ESXi 映像复制到 iSCSI 目标之后，可以执行实际引导。

前提条件

- 将引导网卡上的 iSCSI 引导固件配置为指向引导 LUN。
- 在 BIOS 中更改引导顺序，以便 iSCSI 优先于引导设备。
- 如果使用 Broadcom 适配器，请将**引导至 iSCSI 目标**设置为**已启用**。

步骤

1 重新启动主机。

  主机使用 iBFT 数据从 iSCSI LUN 引导。首次引导过程中，iSCSI 初始化脚本将设置默认网络。在后续重新引导后，网络设置仍持久有效。

2 （可选）使用 vSphere Client 调整网络配置。

网络最佳做法

要使用 iBFT 从 iSCSI 引导 ESXi 主机，必须正确配置网络。

为了获得更好的安全性和更佳性能，主机上需要有冗余网络适配器。

设置所有网络适配器的方式取决于您的环境是使用共享网络，还是使用隔离网络传输 iSCSI 流量和主机管理流量。

共享 iSCSI 网络和管理网络

在主机的第一个网络适配器上配置网络参数和 iSCSI 参数。主机引导后，可以将辅助网络适配器添加到默认的端口组。

隔离 iSCSI 网络和管理网络

配置隔离 iSCSI 网络和管理网络时，请遵循以下准则来避免出现带宽的问题。

- 隔离网络必须位于不同的子网上。
- 如果使用 VLAN 隔离网络，那么它们必须具有不同的子网来确保能够正确设置路由表。
- VMware 建议您将 iSCSI 适配器和 iSCSI 目标配置到相同子网上。如果将 iSCSI 适配器和目标设置到不同的子网上，您必须注意以下限制：
  - 默认的 VMkernel 网关必须能够路由管理流量和 iSCSI 流量。
  - 引导 ESXi 主机后，只能将启用了 iBFT 的网络适配器用于 iBFT。不能将此适配器用于 iSCSI 流量。
- 将第一个物理网络适配器用于管理网络。
- 将第二个物理网络适配器用于 iSCSI 网络。确保配置 iBFT。
- 主机引导后，可以将辅助网络适配器添加到管理网络和 iSCSI 网络。

更改 iBFT iSCSI 引导设置

如果 iSCSI 存储器或主机上的 IQN 名称、IP 地址等设置更改，则需要更新 iBFT。此任务假定引导 LUN 和该 LUN 上存储的数据保持原样。

步骤

1 关闭 ESXi 主机。

2 更改 iSCSI 存储器设置。

3 使用新设置更新主机上的 iBFT。

4 重新启动主机。

  主机使用 iBFT 中存储的新信息进行引导。

对 iBFT iSCSI 引导进行故障排除

本节中的主题帮助您标识和解决使用 iBFT iSCSI 引导时可能会遇到的问题。

系统网关丢失导致网络连接丢失

在删除与 iBFT 网络适配器关联的端口组时丢失网络连接。

问题

删除端口组后丢失网络连接。

原因

安装 ESXi 期间，如果在启用了 iBFT 的网络适配器中指定一个网关，则此网关将成为系统默认网关。如果删除与该网络适配器关联的端口组，系统默认网关将丢失。此操作导致网络连接丢失。

解决方案

除非需要，否则不要设置 iBFT 网关。如果需要此网关，安装后请手动将系统默认网关设置为管理网络使用的网关。

更改 iSCSI 引导参数导致 ESXi 在无状态模式下引导

首次引导后更改网络适配器上的 iSCSI 引导参数不会更新 ESXi 主机上的 iSCSI 和网络配置。

问题

如果在首次从 iSCSI 引导 ESXi 主机后更改了网络适配器上的 iSCSI 引导参数，则主机将在无状态模式下进行引导。

原因

固件使用已更新的引导配置，且能够连接到 iSCSI 目标并加载 ESXi 映像。但加载时，系统不使用新的参数，而是继续使用上次引导时所使用的网络参数和 iSCSI 参数。因此，ESXi 主机无法连接到目标并在无状态模式下引导。

解决方案

1 使用 vSphere Client 连接到 ESXi 主机。

2 重新配置主机上的 iSCSI 参数和网络参数，以与 iBFT 参数相匹配。

3 执行重新扫描。

收集 ESXi 主机的诊断信息

如果使用 iBFT 从 SAN 引导，且 ESXi 主机没有本地诊断分区，那么，您需要设置 vSphere Management Assistant (vMA) 来从主机中收集诊断信息，并存储这些信息以进行分析。

如果 ESXi 主机出现严重问题，它会创建一个文件，其中包含可供日后进行分析以确定问题原因的诊断信息。

可以设置 ESXi 主机，以将其用作在网络上传输诊断文件的网络转储客户端。vMA 则用作网络转储服务器来收集诊断信息，并将这些信息保存到服务器上的 `/var/core` 中的网络转储服务器。

设置网络转储环境时，请遵循以下准则：

- 不要将 vMA 安装在设置网络转储客户端的同一物理主机上。
- 如果有多个 ESXi 主机需要配置网络转储，应分别配置每个主机。一个 vMA 实例足够用来从多个 ESXi 主机中收集核心转储文件。

配置 vSphere Management Assistant

vSphere Management Assistant (vMA) 是一个虚拟机，可对其进行配置使之充当从 ESXi 主机中收集诊断信息的网络转储服务器。

前提条件

在不同于设置为网络转储客户端的物理服务器上部署并配置 vMA。有关信息，请参见 http://www.vmware.com/support/pubs 上的《vSphere Management Assistant 指南》。

步骤

1 以管理员身份登录 vMA。

2 在 vMA 上启用网络转储服务器。

```
# sudo chkconfig vmware-netdumper on
```

3 启动网络转储服务器。

```
# sudo/etc/init.d/vmware-netdumper start

# sudo lokkit --quiet --port=6500:udp
```

配置 ESXi 主机

可以配置 ESXi 主机，以便它充当通过网络传输诊断信息的网络转储客户端。

使用 vSphere CLI 配置 ESXi 主机。

前提条件

将 vMA 配置为网络转储服务器并获取其 IP 地址。

步骤

1 输入 ESXi 主机的 IP 地址。

```
# esxcfg-advcfg -s IP_address_ESXi /Net/NetdumpVmkIP
```

2 指定要使用的网络适配器。

```
# esxcfg-advcfg -s vmnic /Net/NetdumpVmkNic
```

3 指定附加到网络适配器的端口组。

```
# esxcfg-advcfg -s portgroup /Net/NetdumpVmkPG
```

4 指定网络转储服务器的 IP 地址。

```
# esxcfg-advcfg -s IP_address_netdump /Net/NetdumpServerIP
```

5 （可选）输入可访问网络转储服务器的网关的 IP 地址。

```
# esxcfg-advcfg -s IP_address_gateway /Net/NetdumpServerGateway
```

管理使用 SAN 存储器的 ESX/ESXi 系统 5

本节可在管理 ESX/ESXi 系统、有效使用 SAN 存储器以及执行故障排除方面为您提供帮助。另外，本节还将说明如何查找有关存储设备、适配器及多路径等的信息。

本章讨论了以下主题：

查看存储适配器信息

在 vSphere Client 中，可以显示主机使用的存储适配器并检查其信息。

列出所有可用适配器后，可以查看其型号、类型（如光纤通道、并行 SCSI 或 iSCSI）及其唯一标识符（如果可用）。

iSCSI 适配器将 iSCSI 名称用作唯一标识符。

显示每个 iSCSI 适配器的详细信息时，可了解以下信息。需要先配置或启用软件 iSCSI 和从属硬件 iSCSI 适配器后才能查看其信息。

表 5–1 存储适配器信息

适配器信息	描述
型号	适配器的型号。
iSCSI 名称	根据标识 iSCSI 适配器的 iSCSI 标准形成的唯一名称。
iSCSI 别名	用以替代 iSCSI 名称的友好名称。

表 5–1 存储适配器信息 （续）

适配器信息	描述
IP 地址（独立硬件 iSCSI）	分配到 iSCSI 适配器的地址。
已连接的目标	通过适配器访问的目标数。
设备	适配器可以访问的所有存储设备或 LUN。
路径	适配器用于访问存储设备的所有路径。

查看存储适配器信息

使用 vSphere Client 显示存储适配器，并查看其信息。

步骤

1 在“清单”中，选择**主机和群集**。
2 选择主机，然后单击**配置**选项卡。
3 在“硬件”中，选择**存储适配器**。
4 要查看特定适配器的详细信息，请从“存储适配器”列表中选择适配器。
5 要列出适配器可以访问的所有存储设备，请单击**设备**。
6 要列出适配器使用的所有路径，请单击**路径**。

将存储适配器名称复制到剪贴板

可以直接从 UI 将适配器的名称复制到剪贴板。

步骤

1 在“清单”中，选择**主机和群集**。
2 选择主机，然后单击**配置**选项卡。
3 在“硬件”中，选择**存储适配器**。
4 从“存储适配器”列表中选择适配器。
5 在“详细信息”面板中，突出显示名称字段中的值，并从右键单击菜单中选择**复制**。

查看存储设备信息

可以使用 vSphere Client 显示主机可用的所有存储设备或 LUN，包括所有本地和联网的设备。如果使用任何第三方多路径插件，则通过插件可用的存储设备也将出现在列表上。

对于每个存储适配器，可以显示仅通过此适配器可访问的存储设备的单独列表。在检查存储设备的列表时，通常可以看到以下信息。

表 5–2 存储设备信息

设备信息	描述
名称	主机根据存储器类型和制造商为设备分配的友好名称。可以根据需要更改此名称。
标识符	通用唯一标识符是存储设备的固有名称。
运行时名称	设备第一条路径的名称。
LUN	显示目标中 LUN 位置的 LUN 号。
类型	设备类型，例如，磁盘或 CD-ROM。
传输	主机用于访问设备的传输协议。

表 5–2 存储设备信息 （续）

设备信息	描述
容量	存储设备的总容量。
所有者	主机用于管理存储设备的插件（如 NMP 或第三方插件）。
硬件加速	有关存储设备是否帮助主机执行各种虚拟机管理操作的信息。状态可以为“支持”、“不支持”或“未知”。
位置	指向 `/vmfs/devices/` 目录中存储设备的路径。
分区	主分区和逻辑分区，包括 VMFS 数据存储（如果已配置）。

了解存储设备命名

在 vSphere Client 中，每个存储设备或 LUN 均由多个名称进行标识。

名称

主机根据存储器类型和制造商分配给设备的友好名称。您可以使用 vSphere Client 修改名称。

标识符

主机从存储器提取的通用唯一标识符。根据存储器的类型，主机使用不同算法提取标识符。

运行时名称

设备第一条路径的名称。运行时名称由主机创建。此运行时名称不是设备的可靠标识符，并不持久。

运行时名称具有以下格式：

vmhba#:C#:T#:L#，其中

- vmhba# 是存储适配器的名称。此名称指的是主机上的物理适配器，而不是由虚拟机使用的 SCSI 控制器。
- C# 是存储器通道号。

  软件 iSCSI 启动器使用通道号来显示指向同一目标的多个路径。
- T# 是目标号。目标编号由主机决定，可能会在对于主机可见的目标的映射更改时发生变化。由不同主机共享的目标可能没有相同的目标号。
- L# 是显示目标中 LUN 位置的 LUN 号。LUN 号由存储系统提供。如果目标只有一个 LUN，则 LUN 号始终为零 (0)。

例如，`vmhba1:C0:T3:L1` 表示通过存储适配器 vmhba1 和通道 0 访问的目标 3 上的 LUN 1。

显示主机的存储设备

可以使用 vSphere Client 显示主机可用的所有存储设备或 LUN，包括所有本地和联网的设备。如果使用任何第三方多路径插件，则通过插件可用的存储设备也将出现在列表上。

步骤

1 在“清单”中，选择**主机和群集**。

2 选择主机，然后单击**配置**选项卡。

3 在“硬件”中，选择**存储器**。

4 单击**设备**。

5 要查看有关特定设备的其他详细信息，请从列表中选择设备。

显示适配器的存储设备

对于主机上的每个存储适配器，可以显示只有通过此适配器才可访问的存储设备的列表。

步骤

1 在“清单”中，选择**主机和群集**。

2 选择主机，然后单击**配置**选项卡。

3 在“硬件”中，选择**存储适配器**。

4 从“存储适配器”列表中选择适配器。

5 单击**设备**。

将存储设备标识符复制到剪贴板

存储设备标识符是主机分配给存储设备或 LUN 的通用唯一 ID。根据存储器的类型，主机使用不同的算法创建标识符，标识符可以很长很复杂。可以直接从 UI 复制存储设备标识符。

步骤

1 显示存储设备的列表。

2 右键单击设备，然后选择**将标识符复制到剪贴板**。

查看数据存储信息

可以显示对主机可用的所有数据存储，并分析其属性。

可使用以下方式将数据存储添加到 vSphere Client 中：

- 在可用存储设备上创建。
- 当主机添加到清单时发现。将主机添加到清单中时，vSphere Client 将显示对主机可用的任何数据存储。

如果 vSphere Client 连接到 vCenter Server 系统，则可以在“数据存储”视图中查看数据存储信息。此视图按数据中心显示清单中所有的数据存储。通过此视图，可以将数据存储组织到文件夹层次结构中，创建新数据存储，编辑其属性，或移除现有数据存储。

此视图可全面显示数据存储的所有信息，包括使用数据存储的主机和虚拟机、存储报告信息、权限、警报、任务与事件、存储拓扑以及存储报告。“数据存储”视图的“配置”选项卡上提供连接到此数据存储的所有主机上的每个数据存储的配置详细信息。

注意 vSphere Client 直接连接到主机时，“数据存储”视图不可用。在这种情况下，通过主机存储配置选项卡检查数据存储信息。

下表说明查看数据存储时显示的数据存储详细信息。

表 5–3 数据存储信息

数据存储信息	描述
标识	分配给数据存储的可编辑的友好名称。
设备	在其上部署了数据存储的存储设备。如果数据存储跨多个存储设备，则只显示第一个存储设备。
容量	数据存储的总格式化容量。
可用空间	可用空间。
类型	数据存储使用的文件系统，可能是 VMFS 或 NFS。

表 5–3 数据存储信息 （续）

数据存储信息	描述
存储 I/O 控制	允许对群集范围的存储 I/O 进行优化。请参见《vSphere 资源管理指南》。
硬件加速	有关数据存储是否帮助主机执行各种虚拟机管理操作的信息。状态可以为“支持”、“不支持”或“未知”。
位置	指向 `/vmfs/volumes/` 目录中数据存储的路径。
数据区	数据存储跨越的各个数据区及其容量（仅 VMFS 数据存储）。
路径选择	主机用于访问存储（仅限 VMFS 数据存储）的路径选择策略。
路径	用于访问存储及其状态（仅限 VMFS 数据存储）的路径数目。

检查数据存储属性

使用 vSphere Client 检查数据存储属性。

步骤

1 在清单中选择主机，然后单击**配置**选项卡。

2 在“硬件”中，选择**存储器**。

3 单击**数据存储**视图。

4 要显示特定数据存储的详细信息，请从列表中选择数据存储。

解决存储显示问题

当使用 vSphere Client 查看 ESX/ESXi 主机可用的存储设备时，如果输出与您的期望有所不同，请执行故障排除任务。

如果存在显示问题，请执行以下故障排除任务。

表 5–4 排除 iSCSI LUN 显示故障

故障排除任务	描述
检查线缆连接。	如果看不到端口，问题可能在于线缆连接或路由。请先检查线缆。请确保线缆已连接端口，且链路指示灯表示连接良好。如果线缆任意一端的链路指示灯未显示连接良好，请更换线缆。
检查路由设置。	控制以太网配置上的不同子网间的连接。如果 ESX/ESXi 系统和 iSCSI 存储器不在同一个子网上，请确保这些子网间存在适当路由。此外，还需确保在 iSCSI 存储器和 ESX/ESXi 主机中的 iSCSI 启动器上设置了正确的子网掩码和网关地址。
检查访问控制配置。	如果重新扫描后未显示预期的 LUN，则存储系统端可能未正确配置访问控制： ■ 如果配置了 CHAP，请确保 ESX/ESXi 主机上也启用了 CHAP，并与存储系统设置相匹配。 ■ 如果使用基于 IP 的筛选，请确保允许 iSCSI HBA 或 VMkernel 端口组 IP 地址和服务控制台 IP 地址。 ■ 如果使用基于启动器名称的筛选，请确保该名称是符合要求的 iSCSI 名称，并与存储系统设置相匹配。 ■ 要从 SAN 引导，请确保每台主机只会看到所需的 LUN。不要让任何主机看到不归其所有的任何引导 LUN。使用存储系统软件确保主机只能看到应对其可见的 LUN。 ■ 确保通过 **Disk.MaxLUN** 设置可查看预期为可见的 LUN。
检查存储处理器设置。	如果存储系统有多个存储处理器，请确保 SAN 交换机连接了要访问的 LUN 所属的 SP。在某些存储系统上，仅有一个 SP 为主动，而另一 SP 在发生故障之前是被动的。如果连接错误的 SP（对应被动路径的 SP），则您可能看不到预期的 LUN，或可能看到了 LUN，但在尝试进行访问时出错。

表 5–4 排除 iSCSI LUN 显示故障 （续）

故障排除任务	描述
对于软件 iSCSI，请检查网络配置。	ESX/ESXi 中的软件 iSCSI 启动器要求 VMkernel 网络端口对 iSCSI 存储器具有访问权限。该软件启动器使用 VMkernel 在 ESX/ESXi 系统与 iSCSI 存储器之间传输数据。
重新扫描 iSCSI 启动器。	在每次完成以下任务时，执行重新扫描： ■ 在 SAN 上创建新 LUN。 ■ 更改 ESX/ESXi 主机存储系统上的 LUN 屏蔽。 ■ 重新连接线缆。 ■ 对群集中的主机进行更改。 ■ 更改 CHAP 设置或添加新发现地址。

数据存储刷新和存储重新扫描操作

数据存储刷新操作可更新 vSphere Client 中显示的数据存储列表和存储信息（例如数据存储容量）。执行数据存储管理任务或者在 SAN 配置中进行更改时，可能需要重新扫描存储器。

执行 VMFS 数据存储管理操作时，例如创建 VMFS 数据存储或 RDM、添加数据区和增加或删除 VMFS 数据存储，主机或 vCenter Server 将自动重新扫描和更新存储器。可以通过关闭"主机重新扫描筛选器"来禁用自动重新扫描功能。请参见第 70 页，"关闭 vCenter Server 存储筛选器"。

在某些情况下，需要执行手动重新扫描。可以重新扫描您的主机可用的所有存储器，或者如果使用的是 vCenter Server，则可以重新扫描文件夹、群集和数据中心中的所有主机可用的所有存储器。

如果进行的更改只针对通过特定适配器连接的存储器，则需要重新扫描此适配器。

每次进行以下更改之一后请手动执行重新扫描：

- 在 SAN 上创建新 LUN。
- 更改主机上的路径屏蔽。
- 重新连接线缆。
- 更改 CHAP 设置。
- 添加或移除发现或静态地址。
- 在 vCenter Server 中编辑或移除由 vCenter Server 主机和单台主机共享的数据存储之后，向 vCenter Server 中添加该单台主机。

重要事项 如果在某条路径不可用时重新扫描，则主机将从到设备的路径的列表中移除该路径。当该路径再次变得可用并且开始工作时，会立即重新显示在列表中。

执行存储重新扫描

在 SAN 配置中进行更改后，可能需要重新扫描存储器。您可以重新扫描主机可用的所有存储。如果进行的更改只针对通过特定适配器访问的存储器，则只重新扫描此适配器。

如果只需要重新扫描特定主机可用的存储器或通过主机上的特定适配器访问的存储器，请使用此步骤。如果要重新扫描由 vCenter Server 系统管理的所有主机可用的存储，则可以通过右键单击包含这些主机的数据中心、群集或文件夹并选择**重新扫描数据存储**来执行此操作。

步骤

1 在 vSphere Client 中，选择一台主机，然后单击**配置**选项卡。

2 在“硬件”面板中，选择**存储适配器**，然后单击“存储适配器”面板上方的**重新扫描**。

  也可右键单击一个适配器，并单击**重新扫描**只对该适配器进行重新扫描。

  重要事项 在 ESXi 上，不可能重新扫描单个存储适配器。如果重新扫描单个适配器，则所有适配器都会进行重新扫描。

  ---

3 要发现新的磁盘或 LUN，请选择**扫描新的存储设备**。

  新发现的 LUN 将显示在设备列表上。

4 要发现新的数据存储或在其配置更改后更新数据存储，请选择**扫描新的 VMFS 卷**。

  发现的新数据存储或 VMFS 卷将出现在数据存储列表上。

更改扫描的 LUN 数目

默认情况下，VMkernel 从每一目标的 LUN 0 扫描到 LUN 255（总共 256 个 LUN）。可以修改 **Disk.MaxLUN** 参数以提升发现 LUN 的速度。

重要事项 无法发现 LUN ID 号大于 255 的 LUN。

减小该值可缩短重新扫描时间和引导时间。但是，重新扫描 LUN 所需的时间可能取决于其他因素，其中包括存储系统类型以及是否启用了稀疏 LUN 支持。

步骤

1 在 vSphere Client 清单面板中，选择主机，然后依次单击“软件”下的**配置**选项卡和**高级设置**。

2 选择**磁盘**。

3 向下滚动至 **Disk.MaxLUN**。

4 将现有值更改为所选的值，并单击**确定**。

  输入的值会将 LUN 指定为要发现的最后一个 LUN 的后面一个。

  例如，要发现 0 到 31 的 LUN，则将 **Disk.MaxLUN** 设置为 32。

禁用稀疏 LUN 支持

可以禁用默认稀疏 LUN 支持以减少 ESX/ESXi 扫描 LUN 所需的时间。

VMkernel 默认提供稀疏 LUN 支持。稀疏 LUN 支持使 VMkernel 在存储系统显示非顺序编号的 LUN（例如 0、6 和 23）时可以执行不中断的 LUN 扫描。如果存储系统显示的所有 LUN 均为顺序编号，则可以禁用稀疏 LUN 支持。

步骤

1 在 vSphere Client 清单面板中，选择主机，然后依次单击“软件”下的**配置**选项卡和**高级设置**。

2 在“高级设置”对话框中，选择**磁盘**。

3 向下滚动至 **Disk.SupportSparseLUN**，将值更改为 0，然后单击**确定**。

关闭 vCenter Server 存储筛选器

在执行 VMFS 数据存储管理操作时，vCenter Server 使用默认存储筛选器。这些筛选器可以仅检索可用于特定操作的存储设备或 LUN，从而能更有效地避免存储设备损坏。不适合的 LUN 不会显示出来供选择。可以关闭筛选器来查看所有 LUN。

在对 LUN 筛选器进行任何更改之前，请咨询 VMware 支持团队。仅当您有其他方法来防止 LUN 损坏时，才可关闭筛选器。

步骤

1 在 vSphere Client 中，选择**管理 > vCenter Server 设置**。

2 在设置列表中，选择**高级设置**。

3 在**键**文本框中，键入键。

键	筛选器名称
config.vpxd.filter.vmfsFilter	VMFS 筛选器
config.vpxd.filter.rdmFilter	RDM 筛选器
config.vpxd.filter.SameHostAndTransportsFilter	相同主机和传输筛选器
config.vpxd.filter.hostRescanFilter	主机重新扫描筛选器 **注意** 如果关闭“主机重新扫描筛选器”，则每次将一个新的 LUN 提供给主机或群集时，主机仍继续执行重新扫描。

4 在**值**文本框中，为指定的键键入 `False`。

5 单击**添加**。

6 单击**确定**。

无需重新启动 vCenter Server 系统。

vCenter Server 存储筛选

vCenter Server 提供存储筛选器，来帮助您避免由于使用不支持的 LUN 可能导致的存储设备损坏或性能下降。默认情况下，这些筛选器可用。

表 5–5 存储筛选器

筛选器名称	描述	键
VMFS 筛选器	筛选出在 vCenter Server 管理的任何主机上已由 VMFS 数据存储使用的存储设备或 LUN。LUN 不会以要格式化为另一个 VMFS 数据存储或要用作 RDM 的候选者的身份显示。	config.vpxd.filter.vmfsFilter
RDM 筛选器	筛选出在 vCenter Server 管理的任何主机上已由 RDM 引用的 LUN。LUN 不会以要格式化为 VMFS 或要供其他 RDM 使用的候选者的身份显示。 如果需要虚拟机访问相同的 LUN，则虚拟机必须共享相同的 RDM 映射文件。有关此类配置的信息，请参阅《故障切换群集和 Microsoft 群集服务的设置》。	config.vpxd.filter.rdmFilter

表 5–5 存储筛选器 （续）

筛选器名称	描述	键
相同主机和传输筛选器	筛选出由于主机或存储类型不兼容而不适合用作 VMFS 数据存储区的 LUN。防止将以下 LUN 添加为数据区： ■ 未对共享原始 VMFS 数据存储的所有主机公开的 LUN。 ■ 使用不同于原始 VMFS 数据存储使用的存储类型的 LUN。例如，无法将光纤通道数据区添加到本地存储设备上的 VMFS 数据存储。	config.vpxd.filter.SameHostAndTransportsFilter
主机重新扫描筛选器	在执行数据存储管理操作之后，自动重新扫描和更新 VMFS 数据存储。该筛选器可帮助提供 vCenter Server 管理的所有主机上的所有 VMFS 数据存储的一致视图。 **注意** 如果将一个新的 LUN 提供给主机或群集，则无论是否打开了“主机重新扫描筛选器”，主机都会自动执行重新扫描。	config.vpxd.filter.hostRescanFilter

路径扫描和声明

启动 ESX/ESXi 主机或重新扫描存储适配器时，主机会发现它可以使用的存储设备的所有物理路径。根据 `/etc/vmware/esx.conf` 文件中所定义的一组声明规则，主机会确定应声明特定设备路径并负责管理该设备的多路径支持的多路径插件 (MPP)。

默认情况下，主机每 5 分钟执行一次定期路径评估，从而促使相应 MPP 声明任何尚未声明的路径。

对声明规则进行了编号。对于每个物理路径，主机都通过声明规则运行，首先从最小编号开始。然后，会将物理路径的属性与声明规则中的路径规范进行比较。如果二者匹配，主机会分配声明规则中指定的一个 MPP 来管理物理路径。此过程将持续到所有物理路径均由相应 MPP（第三方多路径插件或本机多路径插件 (NMP)）声明后才结束。

有关多路径插件的一般信息，请参见第 18 页，“管理多路径”。

对于由 NMP 模块管理的路径，将应用第二组声明规则。这些规则确定哪些存储阵列类型插件 (SATP) 应当用于管理特定阵列类型的路径，以及哪些路径选择插件 (PSP) 用于各个存储设备。例如，对于属于 EMC CLARiiON CX 存储系列并且未配置为 ALUA 设备的存储设备，默认 SATP 是 VMW_SATP_CX，默认 PSP 是“最近使用”。

使用 vSphere Client 查看主机用于特定存储设备的 SATP 和 PSP，以及该存储设备的所有可用路径的状态。如果需要，可以使用 vSphere Client 更改默认 VMware PSP。要更改默认 SATP，需要使用 vSphere CLI 修改声明规则。

可以在第 93 页，“管理存储路径和多路径插件”中找到有关修改声明规则的部分信息。

有关可用于管理 PSA 的命令的详细说明，请参见《vSphere 命令行界面安装和脚本指南》以及《vSphere 命令行界面参考》。

查看路径信息

使用 vSphere Client 确定 ESX/ESXi 主机用于特定存储设备的 SATP 和 PSP，以及该存储设备的所有可用路径的状态。可以从“数据存储”和“设备”视图访问路径信息。对于数据存储，请检查与部署了数据存储的设备相连的路径。

路径信息包括分配用来管理设备的 SATP、路径选择策略 (PSP)、路径及其物理特性的列表（如适配器和各条路径使用的目标）以及各条路径的状态。其中会显示以下路径状态信息：

活动
可用于对 LUN 发出 I/O 的路径。目前用于传输数据的单个或多个工作路径标记为“活动”(I/O)。

注意 对于运行 ESX/ESXi 3.5 或更低版本的主机，术语“主动”意为只有一条路径被主机用来向 LUN 发出 I/O。

备用
路径处于工作状态，并且在活动路径发生故障时可用于 I/O。

已禁用
路径已禁用，无法传输数据。

失效
软件无法通过此路径连接磁盘。

如果正在使用**固定**路径策略，就可以看到哪一条路径是首选路径。首选路径的“首选”列标有一个星号 (*)。

查看数据存储路径

使用 vSphere Client 检查连接到部署了数据存储的存储设备的路径。

步骤

1 登录 vSphere Client，在“清单”面板中选择服务器。
2 单击**配置**选项卡，然后在“硬件”面板中单击**存储器**。
3 在“查看”下方单击**数据存储**。
4 在已配置的数据存储列表中，选择要查看或配置其路径的数据存储。

   “详细信息”面板显示用来访问设备的路径总数，以及是否有任何路径已中断或已禁用。

5 单击**属性 > 管理路径**以打开“管理路径”对话框。

   可以使用“管理路径”对话框来启用或禁用路径、设置多路径策略，以及指定首选路径。

查看存储设备路径

使用 vSphere Client 查看主机用于特定存储设备的 SATP 和 PSP，以及该存储设备的所有可用路径的状态。

步骤

1 登录 vSphere Client，在“清单”面板中选择服务器。
2 单击**配置**选项卡，然后在“硬件”面板中单击**存储器**。
3 在“查看”下方单击**设备**。
4 单击**管理路径**以打开“管理路径”对话框。

设置路径选择策略

对于每个存储设备，ESX/ESXi 主机都将根据 `/etc/vmware/esx.conf` 文件中所定义的声明规则来设置路径选择策略。

默认情况下，VMware 支持以下路径选择策略。如果在主机上安装了第三方的 PSP，其策略也将显示于列表中。

固定 (VMW_PSP_FIXED)
当通往磁盘的首选路径可用时，主机将始终使用此路径。如果主机无法通过首选路径访问磁盘，它会尝试替代路径。“固定的”是主动-主动存储设备的默认策略。

固定的 AP (VMW_PSP_FIXED_AP)
将“固定”功能扩展到主动-被动阵列和 ALUA 模式阵列。

最近使用 (VMW_PSP_MRU)
主机选择它最近使用的路径。当路径不可用时，主机会选择一个替代路径。当该路径再次可用时，主机不会恢复到原始路径。没有 MRU 策略的首选路径设置。MRU 是主动-被动存储设备的默认策略。

循环 (VMW_PSP_RR)
当连接到主动-被动阵列时，主机使用自动路径选择算法轮流选择所有活动路径；或者，当连接到主动-主动阵列时，主机使用该算法轮流选择所有可用路径。这样可在主机可用的物理路径之间实现负载平衡。

负载平衡是在路径之间分配 I/O 请求的过程。目的是针对吞吐量（例如每秒 I/O 流量、每秒兆字节数或响应时间）实现最佳性能。

表 5-6 总结了主机的行为随不同阵列类型和故障切换策略变化的情况。

表 5–6 路径策略影响

策略/控制器	主动/主动	主动/被动
最近使用	发生路径故障后进行故障恢复需要管理员操作。	发生路径故障后进行故障恢复需要管理员操作。
固定的	连接恢复后，VMkernel 继续使用首选路径。	VMkernel 尝试继续使用首选路径。这会导致路径抖动或故障，因为另一 SP 现拥有 LUN 的所有权。
循环	无故障恢复。	选择了循环调度中的下一路径。
固定的 AP	对于 ALUA 阵列，VMkernel 选择设置为首选路径的路径。 对于 A/A 和 A/P 以及 ALUA 阵列，只有在路径抖动避免算法允许故障恢复时，VMkernel 才继续使用首选路径。	

更改路径选择策略

通常，不需要更改主机用于特定存储设备的默认多路径设置。但是，如果要进行任何更改，可以使用“管理路径”对话框修改路径选择策略并指定“固定”策略的首选路径。

步骤

1 从“数据存储”视图或“设备”视图打开“管理路径”对话框。

2 选择路径选择策略。

   默认情况下，VMware 支持以下路径选择策略。如果在主机上安装了第三方的 PSP，其策略也将显示于列表中。

   - **固定 (VMW_PSP_FIXED)**
   - **固定的 AP (VMW_PSP_FIXED_AP)**
   - **最近使用 (VMW_PSP_MRU)**
   - **循环 (VMW_PSP_RR)**

3 对于“固定”策略，请指定首选路径，方法是：右键单击要作为首选路径分配的路径并选择**首选**。

4 单击**确定**以保存设置并退出对话框。

禁用路径

由于维护或其他原因，可以暂时禁用路径。您可以使用 vSphere Client 完成此操作。

步骤

1 从“数据存储”视图或“设备”视图打开“管理路径”对话框。

2 在“路径”面板中，右键单击要禁用的路径，然后选择**禁用**。

3 单击**确定**以保存设置并退出对话框。

还可以通过右键单击列表中的路径，然后选择**禁用**来从适配器的“路径”视图禁用路径。

路径管理和手动或静态负载平衡

在可用路径间平衡负载将提高性能。如果阵列支持 ALUA 协议，则主机可使用该协议执行负载平衡。如果阵列不支持 ALUA 协议，而您希望主机执行自动负载平衡，请将设备配置为使用“循环”PSP。如果希望亲自执行负载平衡，则可以执行手动负载平衡。

使用主动-主动和主动-被动存储阵列，可以将主机设置为使用不同路径访问不同的 LUN，从而实现适配器的均衡使用。如果某一路径发生故障，其余可用路径将承载所有流量。路径故障切换可能需要一分钟或更长时间，因为 SAN 可能要协同新的拓扑来尝试还原服务。此延迟对于允许 SAN 在拓扑更改后稳定其配置是必要的。

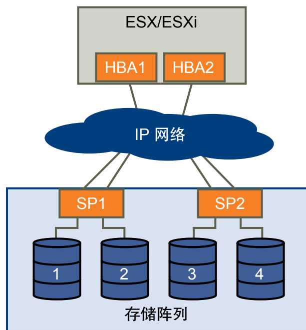
通过将首选的路径分配到 LUN，可以配置 ESX/ESXi 主机以跨多个适配器平衡流量负载。路径策略必须设置为“固定”。

以下示例演示了如何使用主动-主动阵列执行手动负载平衡。

假定以下设置，如图 5‑1 中所示。

- 主动/主动 SP
- 一个 ESX/ESXi 系统
- 两个 iSCSI HBA

图 5–1 手动平衡 iSCSI 的负载



为实现负载平衡，请按如下所示设置首选路径。

- 对于 LUN 1：HBA1-SP1-LUN1
- 对于 LUN 2：HBA2-SP1-LUN2
- 对于 LUN 3：HBA1-SP2-LUN3
- 对于 LUN 4：HBA2-SP2-LUN4

如果阵列可支持两个活动路径，且 HBA 端口可访问阵列中的两个 SP，则可使用主动-被动阵列执行负载平衡。

可使用 VMW_PSP_FIXED_AP 路径选择策略在主动-被动阵列上静态平衡路径负载。

设置客户机操作系统超时

增加标准磁盘超时值，以便 Windows 客户机操作系统不会在路径故障切换过程中发生大范围中断。

路径故障切换在指向 LUN 的活动路径由某一路径更改为另一路径时发生，通常是因为当前路径沿线的某一 SAN 组件发生故障。

I/O 可能会暂停 30 到 60 秒，直到 iSCSI 驱动程序确定该链路不可用并且故障切换结束时为止。因此，虚拟机（其虚拟磁盘安装在 SAN 存储器上）可能变得没有响应。如果尝试显示主机、主机存储设备或适配器，操作可能会停止。故障切换完成后，I/O 恢复正常。

如果存在多处故障，所有与 SAN 存储设备的连接可能都会丢失。如果任何与存储设备的连接均无法工作，某些虚拟机在其虚拟 SCSI 磁盘上可能遇到 I/O 错误。

对于 Windows 2000 和 Windows Server 2003 客户机操作系统，可以使用注册表设置操作系统超时。

步骤

1. 备份 Windows 注册表。
2. 选择**开始 > 运行**。
3. 从命令提示符处键入 **`regedit.exe`**，然后单击**确定**。
4. 在左面板层次结构视图中，依次双击 **HKEY_LOCAL_MACHINE**、**System**、**CurrentControlSet**、**Services** 和 **Disk**。
5. 选中 **TimeOutValue** 并将数据值设置为 x03c（十六进制）或 60（十进制）。

   进行此更改后，Windows 将至少等待 60 秒，以便延迟的磁盘操作完成，然后才会生成错误。
6. 单击**确定**，退出**注册表编辑器**。

共享诊断分区

通常，将 ESX/ESXi 主机的本地磁盘用作诊断分区。但对于从 SAN 引导的无磁盘 ESX 服务器，多个主机可共享同一 SAN LUN 上的一个诊断分区。

如果多个 ESX/ESXi 系统使用一个 LUN 作为诊断分区，必须对该 LUN 进行区域分配以便所有服务器都可对其进行访问。

每台服务器需要 100 MB 空间，因此 LUN 的大小将决定可共享该 LUN 的服务器数目。每个 ESX/ESXi 系统均映射到一个诊断插槽。如果服务器共享诊断分区，VMware 建议磁盘空间至少满足 16 个插槽所需 (1600 MB)。

如果设备上只有一个诊断插槽，所有共享该设备的 ESX/ESXi 系统将映射到同一插槽。这样设置很容易产生问题。如果两个 ESX/ESXi 系统同时执行核心转储，则诊断分区上的核心转储将被覆盖。

避免和解决 SAN 问题

将 ESX/ESXi 与 SAN 一起使用时，必须依照特定准则才能避免 SAN 问题。

应特别留意这些避免和解决 SAN 配置问题的提示：

- 每个 LUN 上仅放置一个 VMFS 数据存储。不建议将多个 VMFS 数据存储放置在同一 LUN 上。
- 不要更改系统为您设置的路径策略，除非您了解做出此类更改的影响。
- 将所有信息记录在案。其中包括涉及以下项目的信息：配置、访问控制、存储器、交换机、服务器和 iSCSI HBA 配置、软件和固件版本以及存储器线缆布局。
- 对故障情况进行规划：
  - 制作多个拓扑映射副本。考虑每一元素发生故障时可能对 SAN 带来的影响。
  - 除去不同链接、交换机、HBA 和其他元素，确保未遗漏设计中的关键故障点。
- 确保根据插槽和总线速度将 iSCSI HBA 安装到 ESX/ESXi 主机中的正确插槽。在服务器中的可用总线之间平衡 PCI 总线负载。
- 在所有可视点（包括 ESX/ESXi 性能图表、以太网交换机统计信息及存储性能统计信息）熟悉存储网络中的不同监控点。
- 更改由 ESX/ESXi 主机使用的 VMFS 数据存储的 LUN 的 ID 时，请务必小心操作。如果您更改此 ID，则 VMFS 数据存储上运行的虚拟机将发生故障。

  如果 VMFS 数据存储上没有正在运行的虚拟机，则在您更改 LUN 的 ID 后，必须使用重新扫描在主机上重置 ID。有关使用重新扫描的详细信息，请参阅第 68 页，“执行存储重新扫描”。

优化 SAN 存储器性能

要优化典型 SAN 环境，需要考虑多个因素。

如果正确配置了网络环境，则 iSCSI 组件将为 iSCSI 启动器和目标提供足够的吞吐量及足够短的延迟。如果网络拥塞，并且链接、交换机或路由器处于饱和状态，则 iSCSI 性能会下降，不能满足 ESX/ESXi 环境的需求。

存储系统性能

存储系统性能是影响整个 iSCSI 环境性能的主要因素之一。

如果存储系统性能出现问题，请参阅存储系统供应商文档以获取任何相关信息。

分配 LUN 时，请记住，您可以通过多台 ESX/ESXi 主机访问每个共享的 LUN，而且每台主机上都可运行多台虚拟机。由 ESX/ESXi 主机使用的一个 LUN 可向运行于不同操作系统的多个不同应用程序提供 I/O 服务。由于此工作负载并非恒定不变，ESX/ESXi LUN 所在的 RAID 组不应包括其他主机所使用的 LUN，这些主机上未运行 ESX/ESXi 用于 I/O 密集型应用程序。

启用读缓存和写缓存。

负载平衡的过程即是将服务器 I/O 请求分散于所有可用 SP 及其关联的主机服务器路径。目的是针对吞吐量（每秒 I/O 流量、每秒兆字节数或响应时间）实现最佳性能。

需要不断对 SAN 存储系统进行重新设计和调试，以确保所有存储系统路径间的 I/O 负载平衡。为满足此要求，请在所有 SP 间分发指向 LUN 的路径以提供最佳负载平衡效果。密切监控可指示何时需要手动重新平衡 LUN 的分发。

调试静态平衡存储系统即是监控特定性能统计信息（例如每秒 I/O 操作数、每秒块数及响应时间）并通过分发 LUN 工作负载将工作负载分散到所有 SP。

注意 ESX/ESXi 目前尚不支持动态负载平衡。

服务器性能

为确保最佳服务器性能，必须考虑几个因素。

各服务器应用程序访问其专用存储器时必须满足以下条件：

- 高 I/O 速率（每秒 I/O 操作数）
- 高吞吐量（每秒兆字节数）
- 最小延迟（响应时间）

由于各应用程序的要求不尽相同，您可以选择存储系统上的适当 RAID 组来实现上述目标。要实现性能目标，请执行以下任务：

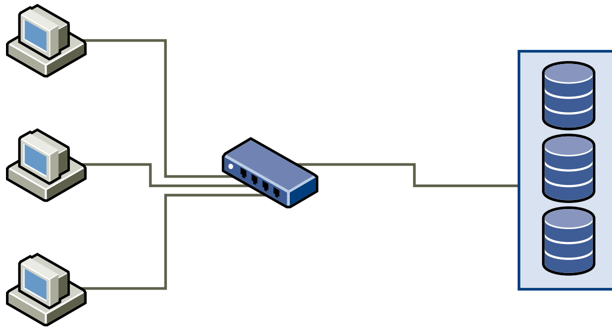
- 将每个 LUN 置于提供必要性能级别的 RAID 组。请注意所分配的 RAID 组中其他 LUN 的活动及资源利用率。对于高性能 RAID 组，因有过多应用程序对其执行 I/O 操作，它可能无法满足 ESX/ESXi 主机上运行的应用程序所需的性能目标。
- 为各服务器提供足够多的网络适配器或 iSCSI 硬件适配器，以便能满足高峰时段服务器上托管的所有应用程序的最大吞吐量。将 I/O 分散到多个端口可为各应用程序提供更高的吞吐量及更短的延迟。
- 要提供软件 iSCSI 的冗余，请确保启动器连接到用于 iSCSI 连接的所有网络适配器。
- 为 ESX/ESXi 系统分配 LUN 或 RAID 组时，多个操作系统将使用和共享该资源。因此，当您使用 ESX/ESXi 系统时，存储子系统中各个 LUN 所需的性能要远高于使用物理机的情况。例如，如果计划运行四个 I/O 密集型应用程序，请为 ESX/ESXi LUN 分配四倍大小的性能容量。
- 将多个 ESX/ESXi 系统与 vCenter Server 一起使用时，存储子系统所需的性能将相应增加。
- ESX/ESXi 系统上运行的应用程序所需的未完成 I/O 数目应与 SAN 能处理的 I/O 数目相匹配。

网络性能

典型的 SAN 包括一组通过交换机网络连接一组存储系统的计算机。多台计算机访问的都是同一个存储器。

图 5‑2 显示了通过以太网交换机连接存储系统的多个计算机系统。在此配置中，每个系统都通过单个以太网链接连接交换机，该交换机也通过单个以太网链接连接存储系统。在具有新式交换机和普通流量的多数配置中，这不是一个问题。

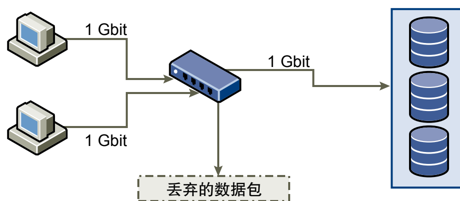
图 5–2 单个以太网链接与存储器的连接



系统从存储器读取数据时，来自存储器的最大响应是发送足够的数据来填充存储系统与以太网交换机之间的链接。任何单个系统或虚拟机都不可能充分利用网络速度，但这是在多个系统共享一个存储设备时可能发生的预期情形。

将数据写入存储器时，多个系统或虚拟机会尝试填充其链接。如图 5–3 所示，发生这种情况时，系统与存储系统间的交换机必须丢弃数据。发生这种情况是因为：使用单个存储设备连接时，要发送至存储系统的流量超过了单个链接所能承载的流量。在这种情况下，交换机会丢弃网络数据包，因为可传输的数据量受到交换机与存储系统间的链接速度的限制。

图 5–3 丢弃的数据包



恢复丢弃的网络数据包会导致性能明显降低。除了花时间确定数据已丢弃，重新传输操作还需使用网络带宽，否则网络带宽将用于当前事务。

iSCSI 流量通过传输控制协议 (Transmission Control Protocol, TCP) 进行网络传输。TCP 是一个可靠的传输协议，可确保重试丢弃的数据包并使其最终到达目标。TCP 旨在恢复并快速、无缝地重新传输丢弃的数据包。但是，如果交换机定期丢弃数据包，网络吞吐量将显著降低。网络会充满重新发送数据的请求以及已重新发送的数据包，并且实际传输的数据量少于网络未拥堵时的数据量。

大多数以太网交换机可缓冲或存储数据，并为每个尝试发送数据的设备提供到达目标的同等机会。这种缓冲部分传输数据的功能与限制未完成命令数目的许多系统相结合，可允许若干系统的小型脉冲依次发送至存储系统。

如果事务较大，且多台服务器正尝试通过单个交换机端口发送数据，则交换机将无法在传输一个请求的同时缓冲另一个请求。在这种情况下，交换机会丢弃无法发送的数据，且存储系统必须请求重新传输丢弃的数据包。例如，如果以太网交换机可在输入端口上缓冲 32 KB，而与其连接的服务器认为可向存储设备发送 256 KB，则会丢弃一些数据。

大多数受管交换机都会提供有关所丢弃数据包的信息，类似以下形式：

```
*: interface is up
IHQ:pkts in input hold queue      IQD:pkts dropped from input queue
OHQ:pkts in output hold queue     OQD:pkts dropped from output queue
RXBS:rx rate (bits/sec)           RXPS:rx rate (pkts/sec)
TXBS:tx rate (bits/sec)           TXPS:tx rate (pkts/sec)
TRTL:throttle count
```

表 5–7 交换机信息示例

接口	IHQ	IQD	OHQ	OQD	RXBS	RXPS	TXBS	TXPS	TRTL
* GigabitEthernet0/1	3	9922	0	0	476303000	62273	477840000	63677	0

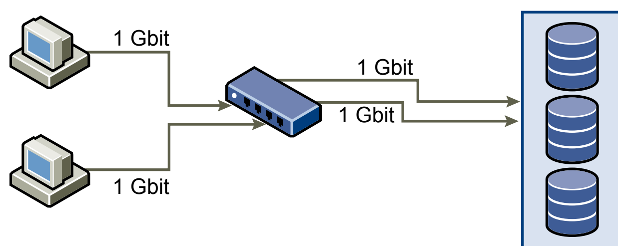
在此 Cisco 交换机示例中，使用的带宽为 476303000 位/秒，低于线速的一半。尽管如此，端口仍缓冲入站数据包，丢弃了许多数据包。此界面摘要最后一行的 IQD 列表示该端口已丢弃近 10,000 个入站数据包。

为避免此问题而对配置进行更改时需确保多个输入以太网链接不缩减为一个输出链接，否则会导致链接超额预定。如果在几乎满载时进行传输的多个链接切换为较小数量的链接，则可能发生超额预定情况。

通常，向存储器写入大量数据的应用程序或系统（例如数据获取或事务日志记录系统）不应共享与存储设备的以太网链接。通过与存储设备的多个连接，这些类型的应用程序可实现最佳性能。

图 5–4 显示了交换机与存储器的多个连接。

图 5–4 交换机与存储器的多个连接



使用 VLAN 或 VPN 不能针对共享配置中的链接超额预定问题提供合适的解决方案。VLAN 和网络的其他虚拟分区提供了逻辑设计网络的方法，但不能更改交换机之间链接和中继的物理功能。当存储流量和其他网络流量最终要共享物理连接时，就像 VPN 一样，可能出现超额预定和丢失数据包的现象。此情况也适用于共享交换机间中继的 VLAN。设计 SAN 的性能时，必须考虑网络的物理限制而不是逻辑分配。

解决性能问题

vSphere Client 提供多种工具用于收集性能信息。这些信息以图形方式显示在 vSphere Client 中。vSphere Client 可定期更新显示屏幕。

还可以使用 `resxtop` vSphere CLI 命令，该命令允许您检查 ESX/ESXi 主机使用资源的方式。有关 `resxtop` 的信息，请参见《资源管理指南》或《vSphere 命令行界面安装和脚本指南》。

检查以太网交换机统计信息

许多以太网交换机提供了监控交换机状况的不同方法。

如果端口经常都需要处理接近最大值的吞吐量，则交换机无法提供最佳性能。如果 iSCSI SAN 中的端口需要处理接近最大值的吞吐量，请减少负载。如果端口与 ESX/ESXi 系统或 iSCSI 存储器连接，则可使用手动负载平衡来减少负载。

如果端口与多个交换机或路由器连接，请考虑在这些组件间安装更多链接以处理更大的负载。以太网交换机通常还提供有关传输错误、排队的数据包和丢弃的以太网数据包的信息。如果交换机定期报告用于 iSCSI 流量的端口上的上述任何状况，iSCSI SAN 的性能将受影响。

解决路径抖动

如果服务器无法访问 LUN，或者访问速度很慢，则可能是路径抖动（也称为 LUN 抖动）的问题。两台主机通过不同 SP 访问 LUN 时，LUN 从未真正可用，此时可能出现路径抖动。

只有特定的 SAN 配置与以下条件同时作用时才会导致路径抖动：

- 使用主动-被动阵列。路径抖动仅发生在主动-被动阵列上。对于主动-主动阵列或提供透明故障切换的阵列，不会出现路径抖动。
- 两个主机使用不同存储处理器 (SP) 访问同一 LUN。可在两种情况下发生路径抖动。
  - 例如，LUN 配置为使用固定 PSP。在主机 A 上，到 LUN 的首选路径设置为使用经由 SP A 的路径。在主机 B 上，到 LUN 的首选路径配置为使用经由 SP B 的路径。
  - 如果主机 A 只能通过经由 SP A 的路径访问 LUN，主机 B 只能通过经由 SP B 的路径访问 LUN，则也可能会发生路径抖动。

直接连接阵列（如 AX100）的一个或多个节点上发生 HBA 故障切换时也可能出现这一问题。

路径抖动问题通常不会出现在其他操作系统中。没有其他常见操作系统使用超过两个服务器共享的 LUN。该设置通常为群集所预留。

如果同时仅一个服务器正在发出到 LUN 的 I/O，则路径抖动不会成为问题。相比而言，多个 ESX/ESXi 系统可能同时向同一 LUN 发出 I/O。

了解路径抖动

在存储阵列中，SP 就像对部分共享存储器具有访问权限的独立计算机。如何处理并发的访问将由算法决定。

对于主动-被动阵列，存储器上组成给定 LUN 的所有扇区每次只能由一个 SP 访问。所有权在存储处理器之间传递。存储系统使用缓存且 SP A 不得向磁盘写入任何使 SP B 缓存无效的内容。由于 SP 在完成操作后必须刷新缓存，因此转移所有权需要一些时间。在此期间，两个 SP 都无法处理 LUN 的 I/O。

对于主动-主动阵列，通过算法可实现对存储器更加精细的访问，并可同步缓存。可通过任何 SP 同时访问，无需额外的时间。

考虑路径选择如何运作：

- 在主动-主动阵列上，ESX/ESXi 系统开始沿新路径发送 I/O。
- 在主动-被动阵列上，ESX/ESXi 系统将检查所有备用路径。当前待考虑路径的 SP 将向系统发送信息，指示其当前是否拥有 LUN 的所有权。
  - 如果 ESX/ESXi 系统找到拥有 LUN 所有权的 SP，则会选定该路径并沿该路径发送 I/O。
  - 如果 ESX/ESXi 主机找不到此类路径，ESX/ESXi 主机便挑选一条备用路径并向 SP 发送命令将 LUN 所有权移至此 SP。

  以下路径选择会导致发生路径抖动：如果服务器 A 只能通过一个 SP 到达某个 LUN，而服务器 B 只能通过另一 SP 到达同一个 LUN，那么二者将使得 LUN 的所有权在两个 SP 之间不断转移，就像打乒乓球。由于系统转移所有权的速度很快，存储阵列无法处理任何 I/O（或只能处理非常少的一部分）。因此，由于 LUN 完成每个 I/O 请求需要花费很长时间，依靠 LUN 的所有服务器都将遭遇吞吐量低下的情形。

解决路径抖动

使用此过程解决路径抖动。两台主机通过不同 SP 访问 LUN 时，LUN 从未真正可用，此时可能在主动-被动阵列上出现路径抖动。

步骤

1 确保在主动-被动阵列上共享同一组 LUN 的所有主机使用同一存储处理器。

2 更正不同 ESX/ESXi 主机与 SAN 目标间的任何线缆接线或屏蔽不一致问题，以便对所有 HBA 显示相同的目标。

3 配置路径以使用“最近使用的 PSP”（默认）或 VMW_PSP_FIXED_AP。

平衡虚拟机之间的磁盘访问

可以在 vSphere Client 中通过 `Disk.SchedNumReqOutstanding` 参数调整未完成的磁盘请求的最大数目。两台或多台虚拟机访问同一 LUN 时，此参数将控制每台虚拟机可向 LUN 发出的未完成请求数量。调整该限制有助于平衡虚拟机之间的磁盘访问。

此限制不适用于 LUN 上只有一个虚拟机处于活动状态的情况。在这种情况下，带宽由存储适配器的队列深度限制。

步骤

1 在 vSphere Client 的“清单”面板中选择主机。

2 单击**配置**选项卡，然后单击“软件”下的**高级设置**。

3 在左侧面板中单击**磁盘**，并向下滚动至 **Disk.SchedNumReqOutstanding**。

4 将参数值更改为所选的数字并单击**确定**。

此更改可对磁盘带宽调度产生影响，但是实验结果证明此更改对磁盘密集型工作负载有改善作用。

下一步

如果在 VMkernel 中调整此值，您可能还需要调整存储适配器中的队列深度。

设置软件 iSCSI 的最大队列深度

如果发现软件 iSCSI LUN 的性能不佳，可通过使用 `vicfg-module` 命令来更改其最大队列深度。

步骤

1 在 vSphere CLI 上，您可运行以下命令：

   `vicfg-module -s iscsivmk_LunQDepth=`*`value`*` iscsi_vmk`

   `iscsivmk_LunQDepth` 参数用于为通过软件 iSCSI 适配器访问的每个 LUN 设置最大未完成命令数或队列深度。默认值为 128。

2 重新引导系统。



小心 将队列深度设置为高于默认值的值，可减少受支持 LUN 的总数。

减少 SCSI 预留

需要在 VMFS 中获取文件锁或元数据锁的操作可导致暂时性 SCSI 预留。SCSI 预留将锁定整个 LUN。某一服务器使用过多 SCSI 预留可导致其他服务器访问相同 VMFS 时性能降低。

需要取得文件锁或元数据锁的操作示例包括：

- 启动虚拟机。
- vMotion。
- 在虚拟机上运行虚拟磁盘快照。
- 需要打开文件或执行元数据更新的文件操作。

注意 仅当 LUN 不支持 VAAI 时，ESX/ESXi 才使用 SCSI 预留机制。如果 LUN 支持 VAAI 且支持硬件加速，则 ESX/ESXi 使用自动测试和设置 (ATS) 算法锁定 LUN。

如果在访问相同 VMFS 的多台服务器上经常发生此类操作，则会导致性能降低。例如，VMware 不建议在相同 VMFS 上运行来自多台服务器的多台使用虚拟磁盘快照的虚拟机。在 VMFS 上运行多台虚拟机时请限制 VMFS 文件操作的数目。

SAN 存储器备份注意事项

在 SAN 环境中，备份有两个目的。第一个目的是将联机数据归档至脱机介质。可对所有联机数据按时间表定期重复执行此过程。第二个目的是提供对脱机数据的访问，用于从故障中恢复。例如，数据库恢复通常需要检索当前未联机的已归档日志文件。

计划备份取决于多种因素：

- 重要应用程序的标识，这些应用程序在给定的一段时间内需要较频繁地备份。
- 恢复点和恢复时间目标。考虑恢复点所需的精确度，以及愿意为此而等待的时间长度。
- 与数据关联的变化率 (Rate of Change, RoC)。例如，如果使用同步/异步复制，RoC 将影响主存储设备与辅助存储设备间所需带宽的大小。
- 对 SAN 环境、存储器性能（备份时）以及其他应用程序的总体影响。
- SAN 上高峰流量时段的标识（计划于这些高峰时段执行的备份会降低应用程序和备份过程的运行速度）。
- 计划数据中心内所有备份的时间。
- 备份单个应用程序所需的时间。
- 归档数据的资源可用性；通常为脱机介质访问（磁带）。

设计备份策略时要包括各应用程序的恢复时间目标。也就是考虑重新置备数据所需的时间和资源。例如，如果已计划的备份要存储过多数据，导致恢复需要大量时间，那么请检查已计划的备份。增加执行备份的频率，这可减少每次备份的数据，从而缩短恢复时间。

如果特定的应用程序需要在某一期限内恢复，则备份过程需要提供时间表及特殊数据处理以满足此需求。快速恢复可能需要使用驻留在联机存储器上的恢复卷，从而尽可能避免通过访问速度较慢的脱机介质来获取缺少的数据组件。

快照软件

管理员通过快照软件可以对磁盘子系统内定义的任一虚拟磁盘生成即时副本。

快照软件分为不同的可用级别：

- 通过 ESX/ESXi 主机可创建虚拟机的快照。此软件在 ESX/ESXi 基本软件包中附送。
- 第三方备份软件可能允许更全面的备份过程，并可能包含更高级的配置选项。

管理员可出于多种原因而生成快照，其中包括：

- 备份
- 灾难恢复
- 多种配置和/或版本的可用性
- 取证 (Forensics)（在系统运行时通过查看快照寻找导致问题的原因）
- 数据挖掘 (Data mining)（查看数据的副本以降低生产系统的负载）

使用第三方备份软件包

使用第三方软件具有环境统一的优点。但是，第三方快照软件带来的额外开销会随着SAN 规模的扩大而变得更高。

如果使用第三方备份软件，请确保 ESX/ESXi 主机支持该软件。

如果使用快照来备份数据，请考虑以下几点：

- 有些供应商同时支持 VMFS 和 RDM 的快照。如果二者均支持，您可以生成主机的整个虚拟机文件系统的快照，也可以生成单个虚拟机的快照（每个磁盘一个快照）。
- 有些供应商仅对使用 RDM 的设置支持快照。如果仅支持 RDM，您可以对单个虚拟机执行快照。

请参见存储供应商的文档。

注意 ESX/ESXi 系统还包括 Consolidated Backup 组件。

分层应用程序

SAN 管理员通常使用基于阵列的专用软件进行备份、灾难恢复、数据挖掘、取证以及配置测试。

存储供应商通常对 LUN 提供两种类型的高级服务：快照和复制。

将 ESX/ESXi 系统与 SAN 一起使用时，您必须确定是基于阵列的工具还是基于主机的工具更适合特定的情形。

基于阵列（第三方）的解决方案

将 ESX/ESXi 系统与 SAN 一起使用时，您必须确定基于阵列的工具是否更适合特定的情形。

考虑基于阵列的解决方案时，请记住以下几点：

- 基于阵列的解决方案通常可得到更全面的统计信息。使用 RDM，数据始终采用同一路径，使得性能管理更为简单。
- 使用 RDM 和基于阵列的解决方案时，安全性对于存储管理员而言更为透明，因为使用 RDM 时虚拟机与物理机更为相似。
- 如果使用基于阵列的解决方案，通常会将物理兼容 RDM 用作虚拟机的存储器。如果不打算使用 RDM，请查看存储供应商的文档，确认是否支持在带有 VMFS 卷的 LUN 上进行操作。如果在 VMFS LUN 上使用阵列操作，请仔细阅读再签名一节。

基于文件 (VMFS) 的解决方案

将 ESX/ESXi 系统与 SAN 一起使用时，您必须确定基于主机的工具是否更适合特定的情形。

考虑使用 VMware Tools 和 VMFS（而非阵列工具）的基于文件的解决方案时，请注意以下几点：

- 使用 VMware Tools 和 VMFS 更利于置备。分配一个较大 LUN，多个 `.vmdk` 文件可置于该 LUN 上。使用 RDM，每台虚拟机都需要一个新 LUN。
- 快照服务随 ESX/ESXi 主机附送，无需额外付费。因此基于文件的解决方案相比基于阵列的解决方案更为经济高效。
- 对 ESX/ESXi 管理员而言，使用 VMFS 更容易。
- ESX/ESXi 管理员若使用基于文件的解决方案可减轻对 SAN 管理员的依赖性。

管理重复 VMFS 数据存储

当 LUN 包含 VMFS 数据存储副本时，您可以使用现有签名或通过分配新签名来挂载该数据存储。

在 LUN 中创建的每个 VMFS 数据存储都有一个唯一的 UUID，该 UUID 存储在文件系统超级块中。对 LUN 进行复制或生成快照后，生成的 LUN 副本的每个字节都与原始 LUN 完全相同。因此，如果原始 LUN 包含具有 UUID X 的 VMFS 数据存储，则 LUN 副本会显示包含具有完全相同 UUID X 的相同的 VMFS 数据存储或 VMFS 数据存储副本。

ESX/ESXi 可以确定 LUN 是否包含 VMFS 数据存储副本，并使用其原始 UUID 挂载数据存储副本，或更改 UUID 从而对该数据存储进行再签名。

使用现有签名挂载 VMFS 数据存储

可能无需再签名 VMFS 数据存储副本。可以挂载 VMFS 数据存储副本，而不更改其签名。

例如，作为灾难恢复计划的一部分，可以在辅助站点上维护虚拟机的同步副本。在主站点发生灾难时，可以在辅助站点上挂载数据存储副本并启动虚拟机。

重要事项 仅当与具有相同 UUID 的原始 VMFS 数据存储不冲突时，才可以挂载 VMFS 数据存储副本。要挂载该副本，原始 VMFS 数据存储必须处于脱机状态。

挂载 VMFS 数据存储时，ESX/ESXi 允许对驻留在 LUN 副本上的数据存储执行读取和写入操作。LUN 副本必须为可写入状态。在系统重新引导后，数据存储挂载也是持久有效的。

由于 ESX/ESXi 不允许对挂载的数据存储进行再签名，所以请在进行再签名之前卸载数据存储。

使用现有签名挂载 VMFS 数据存储

如果不需要对 VMFS 数据存储副本进行再签名，则无需更改其签名即可挂载。

前提条件

在挂载 VMFS 数据存储之前，请在主机上执行存储重新扫描，以便更新为其显示的 LUN 视图。

步骤

1 登录 vSphere Client，然后在“清单”面板中选择服务器。
2 单击**配置**选项卡，然后在“硬件”面板中单击**存储器**。
3 单击**添加存储器**。
4 选择**磁盘/LUN** 存储器类型，然后单击**下一步**。

5 从 LUN 列表中，选择数据存储名称显示在“VMFS 标签”列中的 LUN，然后单击**下一步**。

  “VMFS 标签”列中显示的名称表示 LUN 包含现有 VMFS 数据存储的副本。

6 在“挂载选项”下面，选择**保留现有的签名**。

7 在“即将完成”页面，检查数据存储配置信息，然后单击**完成**。

下一步

如果稍后要对挂载的数据存储进行再签名，则必须先将其卸载。

卸载数据存储

卸载数据存储时，它会保持原样，但是在指定的主机上再也看不到该存储。它继续出现在其他主机上并在这些主机上它保持挂载状态。

只能卸载以下类型的数据存储：

- NFS 数据存储
- 已挂载却没有再签名的 VMFS 数据存储副本

不能卸载活动的已挂载数据存储。

步骤

1 显示数据存储。

2 右键单击要卸载的数据存储，然后选择**卸载**。

3 如果数据存储已共享，请指定不应再访问该数据存储的主机。

  a 如果需要，请取消选中要使其中的数据存储保持挂载状态的主机。

    默认情况下，会选中所有主机。

  b 单击**下一步**。

  c 检查要从中卸载数据存储的主机列表，然后单击**完成**。

4 确认要卸载数据存储。

对 VMFS 副本进行再签名

使用数据存储再签名保留 VMFS 数据存储副本上所存储的数据。对 VMFS 副本进行再签名时，ESX/ESXi 会为副本分配新的 UUID 和新的标签，并将副本挂载为与原始数据存储明显不同的数据存储。

分配到数据存储的新标签的默认格式是 `snap-`*`snapID`*`-`*`oldLabel`*，其中 *snapID* 是整数并且 *oldLabel* 是原始数据存储的标签。

在执行数据存储再签名时，请考虑以下几点：

- 数据存储再签名不可逆。
- 不再将包含再签名的 VMFS 数据存储的 LUN 副本视为 LUN 副本。
- 仅当跨区数据存储的所有数据区联机时，才可对其进行再签名。
- 再签名过程是应急过程，并具有容错性。如果过程中断，可以稍后恢复。
- 可以挂载新的 VMFS 数据存储，而无需承担其 UUID 与其他任何数据存储 UUID 相冲突的风险，如 LUN 快照层次结构中的祖先或子项。

对 VMFS 数据存储副本进行再签名

如果要保留 VMFS 数据存储副本上所存储的数据，请使用数据存储再签名。

前提条件

要对挂载的数据存储副本进行再签名，请先将其卸载。

对 VMFS 数据存储进行再签名之前，请在主机上执行存储重新扫描，以便主机更新为其显示的 LUN 视图并发现所有 LUN 副本。

步骤

1 登录 vSphere Client，然后在“清单”面板中选择服务器。

2 单击**配置**选项卡，然后在“硬件”面板中单击**存储器**。

3 单击**添加存储器**。

4 选择**磁盘/LUN** 存储器类型，然后单击**下一步**。

5 从 LUN 列表中，选择数据存储名称显示在“VMFS 标签”列中的 LUN，然后单击**下一步**。

   “VMFS 标签”列中显示的名称表示 LUN 包含现有 VMFS 数据存储的副本。

6 在“挂载选项”下，选择**分配新签名**，并单击**下一步**。

7 在“即将完成”页面，检查数据存储配置信息，然后单击**完成**。

下一步

进行再签名之后，可能需要执行以下操作：

- 如果再签名的数据存储包含虚拟机，则在虚拟机文件中更新对原始 VMFS 数据存储的引用，这些虚拟机文件包括 `.vmx`、`.vmdk`、`.vmsd` 和 `.vmsn`。
- 要启动虚拟机，请在 vCenter Server 中注册它们。

存储硬件加速

硬件加速功能使主机能够将特定虚拟机和存储管理操作卸载到合规存储硬件。利用存储硬件的辅助，主机可以更快地执行这些操作并且占用更少的 CPU、内存和存储结构带宽。

为实现硬件加速功能，可插入存储架构 (PSA) 结合使用特殊阵列集成插件（称为 VAAI 插件）以及一个阵列集成筛选器（称为 VAAI）。PSA 可自动将 VAAI 筛选器和供应商特定的 VAAI 插件附加到这些支持硬件加速的存储设备。

要查看和管理主机上可用的 VAAI 筛选器和 VAAI 插件，请使用 vSphere CLI 命令。

可以在第 99 页，“管理硬件加速筛选器和插件”中查找有关管理 VAAI 筛选器和 VAAI 插件的一些信息。

有关这些命令的说明，请参见《vSphere 命令行界面安装和脚本指南》和《vSphere 命令行界面参考》。

硬件加速要求和优点

仅当结合使用相应的主机和存储阵列时，硬件加速功能才起作用。

使用下列主机和存储阵列：

- ESX/ESXi 版本 4.1 或更高版本。
- 支持基于存储的硬件加速的存储阵列。ESX/ESXi 版本 4.1 不支持对 NAS 存储设备使用硬件加速。

在主机上，硬件加速功能在默认情况下已启用。要在存储器端启用硬件加速，请咨询存储器供应商。某些存储阵列要求在存储器端明确激活硬件加速支持。

当支持硬件加速功能时，主机在硬件方面可以有所改善，从而可以更快和更有效地执行下列操作：

- 通过 Storage vMotion 迁移虚拟机
- 从模板中部署虚拟机
- 克隆虚拟机或模板
- 虚拟机文件的 VMFS 群集锁定和元数据操作
- 写入精简置备的虚拟磁盘和厚虚拟磁盘
- 创建容错虚拟机

硬件加速支持状态

对于每个存储设备和数据存储，vSphere Client 会在“设备”视图和“数据存储”视图的“硬件加速”列中显示硬件加速支持状态。

状态值为“未知”、“支持”和“不支持”。初始值为“未知”。在主机成功执行卸载操作后，状态将更改为“支持”。如果卸载操作失败，状态将更改为“不支持”。

当存储设备不支持主机操作或只为主机操作提供部分支持时，主机将恢复到其本机方法以执行不支持的操作。

关闭硬件加速

如果存储设备不支持硬件加速功能，可以使用 vSphere Client 高级设置将其关闭。

与任何高级设置一样，在禁用硬件加速之前，请咨询 VMware 支持团队。

步骤

1. 在 vSphere Client 的“清单”面板中，选择主机。
2. 单击**配置**选项卡，然后在**软件**下单击**高级设置**。
3. 单击 VMFS3 并将 **VMFS3.HardwareAcceleratedLocking** 字段中的值更改为零。
4. 单击 **DataMover** 并将以下每个字段中的值更改为零：
   - **DataMover.HardwareAcceleratedMove**
   - **DataMover.HardwareAcceleratedInit**
5. 单击**确定**保存更改。

iSCSI SAN 配置对照表 A

本主题提供不同存储系统和 ESX/ESXi 主机的特殊设置要求对照表。

表 A–1 iSCSI SAN 配置要求

组件	备注
所有存储系统	没有备用电池时必须禁用写入缓存。
拓扑	一处故障不会导致 HBA 和 SP 进行故障切换，尤其是使用主动-被动存储阵列时。
EMC Symmetrix	启用 SPC2 和 SC3 设置。请联系 EMC 获取最新设置。
EMC Clariion	将 EMC Clariion 故障切换模式设置为 1 或 4。有关详细信息，请与 EMC 联系。
HP MSA	无特定要求
HP EVA	对于 EVA3000/5000 固件 4.001 及更高版本和 EVA4000/6000/8000 固件 5.031 及更高版本，将主机类型设置为 `VMware`。 否则将主机模式类型设置为 `Custom`。值为： ■ EVA3000/5000 固件 3.x：000000002200282E ■ EVA4000/6000/8000：000000202200083E
NetApp	如果有任何 iSCSI 启动器是启动器组 (igroup) 的一部分，请针对 NetApp 阵列禁用 ALUA。
EqualLogic	确保已针对硬件 iSCSI 适配器启用 ARP 重定向。
LeftHand	确保已针对硬件 iSCSI 适配器启用 ARP 重定向。
ESX/ESXi 配置	为 ESX/ESXi 主机设置以下高级设置： ■ 将 **Disk.UseLunReset** 设置为 **1** ■ 将 **Disk.UseDeviceReset** 设置为 **0** 对于主动-被动阵列，必须将所有托管群集磁盘的 LUN 的多路径策略设置为“最近使用”。主动-主动阵列上 LUN 的多路径策略可设置为“最近使用”或“固定”。 如果存储系统支持透明故障切换，则允许 ARP 重定向。

VMware vSphere 命令行界面 B

在多数情况下，vSphere Client 非常适合监控连接 SAN 存储器的 ESX/ESXi 主机。高级用户有时可能需要使用某些 VMware vSphere 命令行界面 (vSphere CLI) 命令获取其他详细信息。

有关详细信息，请参见《vSphere 命令行界面安装和脚本指南》。

本附录讨论了以下主题：

resxtop 命令

`resxtop` 命令可实时详细查看 ESX/ESXi 资源使用情况。

有关 `resxtop` 的详细信息，请参见《资源管理指南》和《vSphere 命令行界面安装和脚本指南》。

vicfg-iscsi 命令

`vicfg-iscsi` 命令允许在 ESX/ESXi 主机上配置软件或硬件 iSCSI，设置 CHAP 参数和 iSCSI 网络。

有关详细信息，请参见《vSphere 命令行界面安装和脚本指南》以及《vSphere 命令行界面参考》。

vicfg-mpath 命令

使用 `vicfg-mpath` 命令可查看有关存储设备、路径和多路径插件的信息。

有关详细信息，请参见《vSphere 命令行界面安装和脚本指南》以及《vSphere 命令行界面参考》。

esxcli corestorage claimrule 命令

使用 `esxcli corestorage claimrule` 命令可管理声明规则。声明规则确定哪个多路径模块应当声明特定设备的路径并对该设备进行管理。

有关详细信息，请参见《vSphere 命令行界面安装和脚本指南》以及《vSphere 命令行界面参考》。

vmkping 命令

vmkping 命令允许您验证 VMkernel 网络配置。

用法示例：

```
vmkping [options] [host|IP address]
```

表 B-1 vmkping 选项

选项	描述
-6	使用 IPv6 - ICMPv6 Echo 请求。
-4	使用 IPv4（默认）。
-I	出站接口 - 对于 IPv6 范围。
-D	VMkernel TCP 堆栈调试模式。
-c ***count***	设置数据包计数。
-I ***interval***	设置时间间隔。
-s ***size***	设置发送大小。
-d	在 IPv4 数据包上设置 DF 位。
-v	详细。

管理多路径模块和硬件加速插件 C

使用 vSphere CLI 可管理可插入存储架构 (PSA) 多路径插件和硬件加速插件。

本附录讨论了以下主题：

管理存储路径和多路径插件

通过 vSphere CLI，可以管理 PSA 多路径插件和分配给它们的存储路径。

您可以显示主机上所有可用的多路径插件。可以列出任何第三方 MPP 以及主机的 NMP 和 SATP，并检查它们所声明的路径。还可以定义新路径，并指定应声明路径的多路径插件。

有关可用于管理 PSA 的命令的详细信息，请参见《vSphere 命令行界面安装和脚本指南》以及《vSphere 命令行界面参考》。

列出主机的多路径声明规则

使用 vSphere CLI 可列出可用的多路径声明规则。

声明规则指出给定的物理路径由哪个多路径插件、NMP 或任何第三方 MPP 来管理。每个声明规则基于以下参数标识一组路径：

- 供应商/型号字符串
- 传输（如 SATA、IDE、光纤通道等）
- 适配器、目标或 LUN 位置
- 设备驱动程序（如 Mega-RAID）

步骤

◆ 使用 `esxcli corestorage claimrule list --claimrule-class=MP` 可列出多路径声明规则。

示例 C-1 显示了命令的输出。

示例 C–1 esxcli corestorage claimrule list 命令的示例输出

```
Rule Class  Rule   Class    Type       Plugin     Matches
MP          0      runtime  transport  NMP        transport=usb
MP          1      runtime  transport  NMP        transport=sata
MP          2      runtime  transport  NMP        transport=ide
MP          3      runtime  transport  NMP        transport=block
MP          4      runtime  transport  NMP        transport=unknown
MP          101    runtime  vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          101    file     vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          200    runtime  vendor     MPP_1      vendor=NewVend model=*
MP          200    file     vendor     MPP_1      vendor=NewVend model=*
MP          201    runtime  location   MPP_2      adapter=vmhba41 channel=* target=* lun=*
MP          201    file     location   MPP_2      adapter=vmhba41 channel=* target=* lun=*
MP          202    runtime  driver     MPP_3      driver=megaraid
MP          202    file     driver     MPP_3      driver=megaraid
MP          65535  runtime  vendor     NMP        vendor=* model=*
```

此示例表示以下内容：

- NMP 声明与使用 USB、SATA、IDE 和块 SCSI 传输的存储设备相连的所有路径。
- 默认情况下，MASK_PATH 模块声明所有返回供应商字符串为 DELL 且型号字符串为 Universal Xport 的 SCSI 查询数据的路径。MASK_PATH 用于屏蔽主机的路径。
- MPP_1 模块声明与任何型号 NewVend 存储阵列相连的所有路径。
- MPP_3 模块声明由 Mega-RAID 设备驱动程序控制的存储设备的路径。
- 未在先前的规则中描述的路径由 NMP 声明。
- 输出中的“Rule Class”列说明声明规则的类别。类别可以是 MP（多路径插件）、Filter 或 VAAI。
- “Class”列显示已定义的规则和已加载的规则。“Class”列中的 `file` 参数表示规则已定义。`runtime` 参数表示规则已被加载到系统中。为了激活用户定义的声明规则，应存在两个具有相同规则编号的行，一行是具有 `file` 参数的规则，另一行是具有 `runtime` 的规则。几个低编号的规则只有一行，其“Class”为 `runtime`。这些是系统定义的声明规则，您无法对其进行修改。

显示多路径模块

使用 vSphere CLI 列出已加载到系统中的所有多路径模块。多路径模块管理将主机与存储器相连的物理路径。

步骤

◆ 要列出所有多路径模块，请运行以下命令：

`esxcli corestorage plugin list --plugin-class=MP`

此命令将至少返回 NMP 和 MASK_PATH 模块。如果已加载任何第三方 MPP，也会将它们一并列出。

显示主机的 SATP

使用 vSphere CLI 列出已加载到系统中的所有 VMware NMP SATP。

步骤

◆ 要列出所有 VMware SATP，请运行以下命令：

`esxcli nmp satp list`

对于每个 SATP，该命令将显示如下信息：此 SATP 支持的存储阵列或系统的类型；使用此 SATP 的任何 LUN 的默认 PSP。

请记住以下事项：

- 如果声明规则未将任何 SATP 分配给设备，则 iSCSI 或 FC 设备的默认 SATP 将是 VMW_SATP_DEFAULT_AA。默认 PSP 是 VMW_PSP_FIXED。
- 如果已将 VMW_SATP_ALUA 分配给某特定存储设备，但该设备不能识别 ALUA，则此设备将没有任何匹配的声明规则。在这种情况下，设备将由默认 SATP 根据设备的传输类型进行声明。
- 由 VMW_SATP_ALUA 声明的所有设备的默认 PSP 是 VMW_PSP_MRU。VMW_PSP_MRU 选择由 VMW_SATP_ALUA 报告的主动/优化的路径，如果没有主动/优化的路径，则选择主动/未优化的路径。在有更好的路径可用 (MRU) 之前，将一直使用此路径。例如，如果 VMW_PSP_MRU 当前正在使用主动/未优化的路径，那么，当主动/优化的路径变为可用时，VMW_PSP_MRU 会将当前路径切换到主动/优化的路径。

显示 NMP 存储设备

使用 vSphere CLI 列出由 VMware NMP 控制的所有存储设备，并显示与每个设备关联的 SATP 和 PSP 信息。

步骤

1 要列出所有存储设备，请运行以下命令：

   esxcli nmp device list

2 要显示特定设备的信息，请运行以下命令：

   esxcli nmp device list -d *device_ID*

添加多路径声明规则

使用 vSphere CLI 将新的多路径 PSA 声明规则添加到系统上的一组声明规则。为了激活新声明规则，请先定义规则，然后将其加载到系统中。

例如，当加载新的多路径插件 (MPP) 并需要定义此模块应当声明的路径时，需要添加新的 PSA 声明规则。如果添加新路径并需要用现有的 MPP 对它们进行声明，则可能需要创建一条声明规则。



小心 在创建新声明规则时，请注意避免出现以下情况：相同 LUN 的不同物理路径由不同的 MPP 进行声明。除非其中某个 MPP 是 MASK_PATH MPP，否则，此配置将导致出现性能问题。

步骤

1 要定义新声明规则，请在 vSphere CLI 上运行以下命令：

   esxcli corestorage claimrule add

   有关命令所需选项的信息，请参阅第 101 页，“esxcli corestorage claimrule 选项”。

2 要将新声明规则加载到系统中，请运行以下命令：

   esxcli corestorage claimrule load

   此命令会从系统的配置文件中加载所有新创建的多路径声明规则。

示例 C-2 定义多路径声明规则

- 添加编号为 500 的规则，为 NMP 插件声明型号字符串为 NewMod 以及供应商字符串为 NewVend 的所有路径。

```
# esxcli corestorage claimrule add -r 500 -t vendor -V NewVend -M NewMod -P NMP
```

加载声明规则并运行 **esxcli corestorage claimrule list** 命令之后，可以看到新的声明规则出现在列表中。

注意 声明规则的两行，一行中的“Class”为 runtime，另一行中的“Class”为 file，表示新声明规则已加载到系统中且处于活动状态。

```
Rule Class  Rule   Class    Type       Plugin     Matches
MP          0      runtime  transport  NMP        transport=usb
MP          1      runtime  transport  NMP        transport=sata
MP          2      runtime  transport  NMP        transport=ide
MP          3      runtime  transport  NMP        transport=block
MP          4      runtime  transport  NMP        transport=unknown
MP          101    runtime  vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          101    file     vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          500    runtime  vendor     NMP        vendor=NewVend model=NewMod
MP          500    file     vendor     NMP        vendor=NewVend model=NewMod
```

- 添加编号为 321 的规则，为 NMP 插件声明适配器 vmhba0、通道 0、目标 0 和 LUN 0 上的路径。

```
# esxcli corestorage claimrule add -r 321 -t location -A vmhba0 -C 0 -T 0 -L 0 -P NMP
```

- 添加编号为 1015 的规则，为 NMP 插件声明由光纤通道适配器提供的所有路径。

```
# esxcli corestorage claimrule add -r 1015 -t transport -R fc -P NMP
```

- 添加具有系统分配的规则 ID 的规则，为 NMP 插件声明由光纤通道类型的适配器提供的所有路径。

```
# esxcli corestorage claimrule add --autoassign -t transport -R fc -P NMP
```

删除多路径声明规则

使用 vSphere CLI 将多路径 PSA 声明规则从系统上的一组声明规则中移除。

步骤

1 从声明规则集中删除声明规则。

```
esxcli corestorage claimrule delete -r claimrule_ID
```

有关命令使用的选项的信息，请参阅第 101 页，“esxcli corestorage claimrule 选项”。

注意 默认情况下，PSA 声明规则 101 会屏蔽 Dell 阵列伪设备。除非您要取消屏蔽这些设备，否则请不要删除此规则。

2 从 ESX/ESXi 系统移除声明规则。

```
esxcli corestorage claimrule load
```

屏蔽路径

可以阻止 ESX/ESXi 主机访问存储设备或 LUN，或阻止其使用指向某个 LUN 的单个路径。使用 vSphere CLI 命令可屏蔽路径。

屏蔽路径时，请创建声明规则，该规则将 MASK_PATH 插件分配给指定路径。

步骤

1 检查下一个可用的规则 ID 是多少。

```
esxcli corestorage claimrule list
```

用于屏蔽路径的声明规则的规则 ID 都应该在 101 – 200 范围内。如果此命令显示规则 101 和 102 已经存在，则可以指定 103 来添加规则。

2 通过为 MASK_PATH 插件创建新声明规则，将该插件分配给某个路径。

```
esxcli corestorage claimrule add -P MASK_PATH
```

有关命令行选项的信息，请参见第 101 页，“esxcli corestorage claimrule 选项”。

3 将 MASK_PATH 声明规则加载到系统中。

```
esxcli corestorage claimrule load
```

4 验证是否正确添加了 MASK_PATH 声明规则。

```
esxcli corestorage claimrule list
```

5 如果存在已屏蔽路径的声明规则，则将规则移除。

```
esxcli corestorage claiming unclaim
```

6 运行路径声明规则。

```
esxcli corestorage claimrule run
```

在将 MASK_PATH 插件分配给路径之后，路径状态变为不相关，且不再由主机进行维护。因此，显示屏蔽路径信息的命令可能将路径状态显示为失效。

示例 C-3 屏蔽 LUN

在此示例中，屏蔽目标 T1 和 T2 上通过存储适配器 vmhba2 和 vmhba3 访问的 LUN 20。

1 `#esxcli corestorage claimrule list`

2
```
#esxcli corestorage claimrule add -P MASK_PATH -r 109 -t location -A vmhba2 -C 0 -T 1 -L 20
#esxcli corestorage claimrule add -P MASK_PATH -r 110 -t location -A vmhba3 -C 0 -T 1 -L 20
#esxcli corestorage claimrule add -P MASK_PATH -r 111 -t location -A vmhba2 -C 0 -T 2 -L 20
#esxcli corestorage claimrule add -P MASK_PATH -r 112 -t location -A vmhba3 -C 0 -T 2 -L 20
```

3 `#esxcli corestorage claimrule load`

4 `#esxcli corestorage claimrule list`

5
```
#esxcli corestorage claiming unclaim -t location -A vmhba2
#esxcli corestorage claiming unclaim -t location -A vmhba3
```

6 `# esxcli corestorage claimrule run`

取消路径屏蔽

当需要主机访问被屏蔽的存储设备时，请取消对该设备路径的屏蔽。

步骤

1 删除 MASK_PATH 声明规则。

```
esxcli conn_options corestorage claimrule delete -r rule#
```

2 验证是否正确删除了此声明规则。

```
esxcli conn_options corestorage claimrule list
```

3 将配置文件中的路径声明规则重新加载到 VMkernel 中。

esxcli *conn_options* corestorage claimrule load

4 对指向被屏蔽的存储设备的每条路径运行 `esxcli corestorage claiming unclaim` 命令。

例如：

esxcli *conn_options* corestorage claiming unclaim -t location -A vmhba0 -C 0 -T 0 -L 149

5 运行路径声明规则。

esxcli *conn_options* corestorage claimrule run

您的主机即可访问之前被屏蔽的存储设备。

定义 NMP SATP 规则

NMP SATP 声明规则指定特定的存储设备应当由哪些 SATP 来管理。通常不需要修改 NMP SATP 规则。如果需要这样做，请使用 vSphere CLI 将规则添加到指定 SATP 的声明规则列表中。

当安装特定存储阵列的第三方 SATP 时，可能需要创建一条 SATP 规则。

步骤

1 要添加特定 SATP 的声明规则，请运行 `esxcli nmp satp addrule` 命令。该命令采用以下选项。

选项	描述
-c|--claim-option	添加 SATP 声明规则时设置声明选项字符串。当 SATP 声明路径时，此字符串将传递到 SATP。此字符串的内容和由此导致的 SATP 的行为方式，对每个 SATP 都是唯一的。例如，某些 SATP 支持声明选项字符串 `tpgs_on` 和 `tpgs_off`。如果指定了 `tpgs_on`，则仅当在存储设备上启用了 ALUA 目标端口组支持时，SATP 才将声明路径。
-e|--description	添加 SATP 声明规则时设置声明规则描述。
-d|--device	添加 SATP 声明规则时设置设备。设备规则与供应商/型号规则和驱动程序规则互斥。
-D|--driver	添加 SATP 声明规则时设置驱动程序字符串。驱动程序规则与供应商/型号规则互斥。
-f|--force	强制声明规则在任何情况下均忽略有效性检查并安装规则。
-h|--help	显示帮助消息。
-M|--model	添加SATP声明规则时设置型号字符串。供应商/型号规则与驱动程序规则互斥。
-o|--option	添加 SATP 声明规则时设置选项字符串。
-P|--psp	设置 SATP 声明规则的默认 PSP。
-O|--psp-option	设置 SATP 声明规则的 PSP 选项。
-s|--satp	将为其添加新规则的 SATP。
-R|--transport	添加 SATP 声明规则时设置声明传输类型字符串。
-V|--vendor	添加 SATP 声明规则时设置供应商字符串。供应商/型号规则与驱动程序规则互斥。

注意 当搜索 SATP 规则以查找给定设备的 SATP 时，NMP 会首先搜索驱动程序规则。如果没有匹配项，则搜索供应商/型号规则，最后搜索传输规则。如果仍然没有匹配项，NMP 将选择设备的默认 SATP。

2 要从指定 SATP 的声明规则列表中删除规则，请运行以下命令。

在运行此命令时可以使用与 addrule 相同的选项。

esxcli nmp satp deleterule

3 重新引导主机。

示例 C–4 定义 NMP SATP 规则

以下示例命令将分配 VMW_SATP_INV 插件以管理供应商字符串为 NewVend 和型号字符串为 NewMod 的存储阵列。

```
# esxcli nmp satp addrule -V NewVend -M NewMod -s VMW_SATP_INV
```

如果运行 **esxcli nmp satp listrules -s VMW_SATP_INV** 命令，则可以看到新规则已添加到 VMW_SATP_INV 规则列表中。

管理硬件加速筛选器和插件

硬件加速（或VAAI）筛选器与特定于供应商的VAAI插件一起被附加到支持硬件加速的存储设备。通过 vSphere CLI，可以显示和操作 VAAI 筛选器以及 VAAI 插件。

显示硬件加速筛选器

使用 vSphere CLI 可查看当前加载到系统中的硬件加速（或 VAAI）筛选器。

步骤

◆ 运行 `esxcli corestorage plugin list --plugin-class=Filter` 命令。

此命令的输出与以下类似：

```
Plugin name  Plugin class
VAAI_FILTER  Filter
```

显示硬件加速插件

使用 vSphere CLI 可查看当前加载到系统中的硬件加速插件（也称为 VAAI 插件）。

步骤

◆ 运行 `esxcli corestorage plugin list --plugin-class=VAAI` 命令。

此命令的输出与以下类似：

```
Plugin name         Plugin class
VMW_VAAIP_EQL       VAAI
VMW_VAAIP_NETAPP    VAAI
VMW_VAAIP_CX        VAAI
```

验证存储设备的硬件加速状态

使用 vSphere CLI 可验证特定存储设备的硬件加速支持状态。此命令还有助于确定附加到设备的 VAAI 筛选器。

步骤

◆ 运行 **esxcli corestorage device list --d *device_ID*** 命令。

输出显示硬件加速（或 VAAI）状态，该状态可以为未知、受支持或不受支持。如果设备支持硬件加速，则输出还会列出附加到该设备的 VAAI 筛选器。

```
# esxcli corestorage device list --d naa.60a98000572d43595a4a52644473374c
naa.60a98000572d43595a4a52644473374c
Display Name:NETAPP Fibre Channel Disk(naa.60a98000572d43595a4a52644473374c)
Size: 20480
Device Type:Direct-Access
Multipath Plugin:NMP
Devfs Path:/vmfs/devices/disks/naa.60a98000572d43595a4a52644473374c
Vendor:NETAPP
```

```
Model:LUN
Revision: 8000
SCSI Level: 4
Is Pseudo:false
Status:on
Is RDM Capable:true
Is Local:false
Is Removable:false
Attached Filters:VAAI_FILTER
VAAI Status:supported
Other UIDs:vml.020003000060a98000572d43595a4a52644473374c4c554e202020
```

查看设备的硬件加速插件

使用 vSphere CLI 可查看附加到支持硬件加速的存储设备的硬件加速（或 VAAI）插件。

步骤

◆ 运行 **esxcli vaai device list --d *device_ID*** 命令。

例如：

```
# esxcli vaai device list -d naa.6090a028d00086b5d0a4c44ac672a233
naa.6090a028d00086b5d0a4c44ac672a233
Device Display Name:EQLOGIC iSCSI Disk (naa.6090a028d00086b5d0a4c44ac672a233)
VAAI Plugin Name:VMW_VAAIP_EQL
```

列出硬件加速声明规则

对于支持硬件加速功能的每个存储设备，声明规则会指定硬件加速筛选器和硬件加速插件来管理此存储设备。可以使用 vSphere CLI 列出硬件加速筛选器和插件声明规则。

步骤

1 要列出筛选器声明规则，请运行 esxcli corestorage claimrule list --claimrule-class=Filter 命令。

在此示例中，筛选器声明规则指定应由 VAAI_FILTER 筛选器声明的设备。

```
# esxcli corestorage claimrule list --claimrule-class=Filter
Rule Class   Rule   Class   Type   Plugin      Matches
Filter       65430  runtime vendor VAAI_FILTER vendor=EMC model=SYMMETRIX
Filter       65430  file    vendor VAAI_FILTER vendor=EMC model=SYMMETRIX
Filter       65431  runtime vendor VAAI_FILTER vendor=DGC model=*
Filter       65431  file    vendor VAAI_FILTER vendor=DGC model=*
```

2 要列出 VAAI 插件声明规则，请运行 esxcli corestorage claimrule list --claimrule-class=VAAI 命令。

在此示例中，VAAI 声明规则指定应由特定 VAAI 插件声明的设备。

```
esxcli corestorage claimrule list --claimrule-class=VAAI
Rule Class   Rule   Class   Type    Plugin         Matches
VAAI         65430  runtime vendor  VMW_VAAIP_SYMM vendor=EMC model=SYMMETRIX
VAAI         65430  file    vendor  VMW_VAAIP_SYMM vendor=EMC model=SYMMETRIX
VAAI         65431  runtime vendor  VMW_VAAIP_CX   vendor=DGC model=*
VAAI         65431  file    vendor  VMW_VAAIP_CX   vendor=DGC model=*
```

添加硬件加速声明规则

要为新阵列配置硬件加速，需要添加两个声明规则，一个用于 VAAI 筛选器，而另一个用于 VAAI 插件。要激活新声明规则，首先需要定义这些规则，然后将它们加载到系统中。

步骤

1. 使用 `esxcli corestorage claimrule add --claimrule-class=Filter --plugin=VAAI_FILTER` 命令，为 VAAI 筛选器定义新声明规则。

   有关命令所需选项的信息，请参阅第 101 页，“esxcli corestorage claimrule 选项”。

2. 使用 `esxcli corestorage claimrule add --claimrule-class=VAAI` 命令，为 VAAI 插件定义新声明规则。
3. 使用以下命令加载这两个声明规则：

   ```
   esxcli corestorage claimrule load --claimrule-class=Filter
   esxcli corestorage claimrule load --claimrule-class=VAAI
   ```

4. 使用 `esxcli corestorage claimrule run --claimrule-class=Filter` 命令，运行 VAAI 筛选器声明规则。

   注意 只需要运行 Filter 类规则。当 VAAI 筛选器声明设备时，会自动查找要附加的适当 VAAI 插件。

示例 C-5 定义硬件加速声明规则

要为使用 VMW_VAAI_T10 插件的 IBM 阵列配置硬件加速，请使用以下命令序列：

```
# esxcli corestorage claimrule add --claimrule-class=Filter --plugin=VAAI_FILTER --type=vendor --vendor=IBM --autoassign
# esxcli corestorage claimrule add --claimrule-class=VAAI --plugin=VMW_VAAI_T10 --type=vendor --vendor=IBM --autoassign
# esxcli corestorage claimrule load --claimrule-class=Filter
# esxcli corestorage claimrule load --claimrule-class=VAAI
# esxcli corestorage claimrule run --claimrule-class=Filter
```

删除硬件加速声明规则

使用 vSphere CLI 可删除现有硬件加速声明规则。

步骤

- 您可使用以下命令：

  ```
  esxcli corestorage claimrule delete -r claimrule_ID --claimrule-class=Filter
  esxcli corestorage claimrule delete -r claimrule_ID --claimrule-class=VAAI
  ```

esxcli corestorage claimrule 选项

某些 `esxcli corestorage claimrule` 命令（例如为添加新声明规则、移除规则或屏蔽路径而运行的命令）需要您指定多个选项。

表 C-1 esxcli corestorage claimrule 选项

选项	描述
-A|--adapter	表示此操作中要使用的路径的适配器。
-u|--autoassign	系统将自动分配规则 ID。

表 C‒1 esxcli corestorage claimrule 选项 （续）

选项	描述
-C|--channel	表示此操作中要使用的路径的通道。
-c|--claimrule-class	表示此操作中要使用的声明规则类。 有效值为：MP、Filter 和 VAAI
-d|--device	表示要用于此操作的设备 UID。
-D|--driver	表示此操作中要使用的路径的驱动程序。
-f|--force	强制声明规则在任何情况下均忽略有效性检查并安装规则。
-h|--help	显示帮助消息。
-L|--lun	表示此操作中要使用的路径的 LUN。
-M|--model	表示此操作中要使用的路径的型号。
-P|--plugin	表示要用于此操作的 PSA 插件。
-r|--rule	表示要用于此操作的声明规则 ID。
-T|--target	表示此操作中要使用的路径的目标。
-R|--transport	表示此操作中要使用的路径的传输。 有效值为：block、fc、iscsi、iscsivendor、ide、sas、sata、usb、parallel 和 unknown。
-t|--type	表示用于声明/取消声明或声明规则的匹配类型。 有效值为：vendor、location、driver、transport 和 device。
-V|--vendor	表示要在此操作中使用的路径的供应商。

索引

P

Q

R

S

T

V

W